(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
20 December 2001 (20.12.2001)

PCT

(10) International Publication Number
WO 01/96389 A2

(51) International Patent Classification[7]: C07K 14/47

(21) International Application Number: PCT/US01/18574

(22) International Filing Date: 7 June 2001 (07.06.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/210,667 9 June 2000 (09.06.2000) US
60/252,614 22 November 2000 (22.11.2000) US

(71) Applicant *(for all designated States except US)*: **CORIXA CORPORATION** [US/US]; 1124 Columbia Street, Suite 200, Seattle, WA 98104 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **MEAGHER, Madeleine, Joy** [US/US]; 507 N.E. 71st, #1, Seattle, WA 98115 (US). **KING, Gordon, E.** [US/US]; 1530 N.W. 52nd, #304, Seattle, WA 98107 (US). **XU, Jiangchun** [US/US]; 15805 S.E. 43rd Place, Bellevue, WA 98006 (US). **SECRIST, Heather** [US/US]; 3844 35th Avenue W., Seattle, WA 98199 (US).

(74) Agents: **POTTER, Jane, E., R.**; Seed Intellectual Property Law Group PLLC, Suite 6300, 701 Fifth Avenue, Seattle, WA 98104-7092 et al. (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:
— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 01/96389 A2

(54) Title: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS OF COLON CANCER

(57) Abstract: Compositions and methods for the therapy and diagnosis of cancer, such as colon cancer, are disclosed. Compositions may comprise one or more colon tumor proteins, immunogenic portions thereof, or polynucleotides that encode such portions. Alternatively, a therapeutic composition may comprise an antigen presenting cell that expresses a colon tumor protein, or a T cell that is specific for cells expressing such a protein. Such compositions may be used, for example, for the prevention and treatment of diseases such as colon cancer. Diagnostic methods based on detecting a colon tumor protein, or mRNA encoding such a protein, in a sample are also provided.

COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS OF COLON CANCER

TECHNICAL FIELD OF THE INVENTION

The present invention relates generally to therapy and diagnosis of cancer, such as colon cancer. The invention is more specifically related to polypeptides comprising at least a portion of a colon tumor protein, and to polynucleotides encoding such polypeptides. Such polypeptides and polynucleotides may be used in vaccines and pharmaceutical compositions for prevention and treatment of colon cancer and for the diagnosis and monitoring of such cancers.

BACKGROUND OF THE INVENTION

Cancer is a significant health problem throughout the world. Although advances have been made in detection and therapy of cancer, no vaccine or other universally successful method for prevention or treatment is currently available. Current therapies, which are generally based on a combination of chemotherapy or surgery and radiation, continue to prove inadequate in many patients.

Colon cancer is the second most frequently diagnosed malignancy in the United States as well as the second most common cause of cancer death. The five-year survival rate for patients with colorectal cancer detected in an early localized stage is 92%; unfortunately, only 37% of colorectal cancer is diagnosed at this stage. The survival rate drops to 64% if the cancer is allowed to spread to adjacent organs or lymph nodes, and to 7% in patients with distant metastases.

The prognosis of colon cancer is directly related to the degree of penetration of the tumor through the bowel wall and the presence or absence of nodal involvement, consequently, early detection and treatment are especially important. Currently, diagnosis is aided by the use of screening assays for fecal occult blood, sigmoidoscopy, colonoscopy and double contrast barium enemas. Treatment regimens are determined by the type and stage of the cancer, and include surgery, radiation therapy and/or chemotherapy. Recurrence following surgery (the most common form of therapy) is a major problem and is often the ultimate cause of death. In spite of

considerable research into therapies for the disease, colon cancer remains difficult to diagnose and treat. In spite of considerable research into therapies for these and other cancers, colon cancer remains difficult to diagnose and treat effectively. Accordingly, there is a need in the art for improved methods for detecting and treating such cancers. The present invention fulfills these needs and further provides other related advantages.

SUMMARY OF THE INVENTION

In one aspect, the present invention provides polynucleotide compositions comprising a sequence selected from the group consisting of:

(a) sequences provided in SEQ ID NO:1-377;

(b) complements of the sequences provided in SEQ ID NO:1-377;

(c) sequences consisting of at least 20, 25, 30, 35, 40, 45, 50, 75 and 100 contiguous residues of a sequence provided in SEQ ID NO:1-377;

(d) sequences that hybridize to a sequence provided in SEQ ID NO:1-377, under moderate or highly stringent conditions;

(e) sequences having at least 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98% or 99% identity to a sequence of SEQ ID NO:1-377;

(f) degenerate variants of a sequence provided in SEQ ID NO:1-377.

In one preferred embodiment, the polynucleotide compositions of the invention are expressed in at least about 20%, more preferably in at least about 30%, and most preferably in at least about 50% of colon tumors samples tested, at a level that is at least about 2-fold, preferably at least about 5-fold, and most preferably at least about 10-fold higher than that for normal tissues.

The present invention, in another aspect, provides polypeptide compositions comprising an amino acid sequence that is encoded by a polynucleotide sequence described above.

In certain preferred embodiments, the polypeptides and/or polynucleotides of the present invention are immunogenic, *i.e.*, they are capable of eliciting an immune response, particularly a humoral and/or cellular immune response, as further described herein.

The present invention further provides fragments, variants and/or derivatives of the disclosed polypeptide and/or polynucleotide sequences, wherein the fragments, variants and/or derivatives preferably have a level of immunogenic activity of at least about 50%, preferably at least about 70% and more preferably at least about 90% of the level of immunogenic activity of a polypeptide sequence encoded by a polynucleotide sequence set forth in SEQ ID NO:1-377.

The present invention further provides polynucleotides that encode a polypeptide described above, expression vectors comprising such polynucleotides and host cells transformed or transfected with such expression vectors.

Within other aspects, the present invention provides pharmaceutical compositions comprising a polypeptide or polynucleotide as described above and a physiologically acceptable carrier.

Within a related aspect of the present invention, the pharmaceutical compositions, *e.g.*, vaccine compositions, are provided for prophylactic or therapeutic applications. Such compositions generally comprise an immunogenic polypeptide or polynucleotide of the invention and an immunostimulant, such as an adjuvant.

The present invention further provides pharmaceutical compositions that comprise: (a) an antibody or antigen-binding fragment thereof that specifically binds to a polypeptide of the present invention, or a fragment thereof; and (b) a physiologically acceptable carrier.

Within further aspects, the present invention provides pharmaceutical compositions comprising: (a) an antigen presenting cell that expresses a polypeptide as described above and (b) a pharmaceutically acceptable carrier or excipient. Illustrative antigen presenting cells include dendritic cells, macrophages, monocytes, fibroblasts and B cells.

Within related aspects, pharmaceutical compositions are provided that comprise: (a) an antigen presenting cell that expresses a polypeptide as described above and (b) an immunostimulant.

The present invention further provides, in other aspects, fusion proteins that comprise at least one polypeptide as described above, as well as polynucleotides encoding such fusion proteins, typically in the form of pharmaceutical compositions,

e.g., vaccine compositions, comprising a physiologically acceptable carrier and/or an immunostimulant. The fusions proteins may comprise multiple immunogenic polypeptides or portions/variants thereof, as described herein, and may further comprise one or more polypeptide segments for facilitating the expression, purification and/or immunogenicity of the polypeptide(s).

Within further aspects, the present invention provides methods for stimulating an immune response in a patient, preferably a T cell response in a human patient, comprising administering a pharmaceutical composition described herein. The patient may be afflicted with colon cancer, in which case the methods provide treatment for the disease, or patient considered at risk for such a disease may be treated prophylactically.

Within further aspects, the present invention provides methods for inhibiting the development of a cancer in a patient, comprising administering to a patient a pharmaceutical composition as recited above. The patient may be afflicted with colon cancer, in which case the methods provide treatment for the disease, or patient considered at risk for such a disease may be treated prophylactically.

The present invention further provides, within other aspects, methods for removing tumor cells from a biological sample, comprising contacting a biological sample with T cells that specifically react with a polypeptide of the present invention, wherein the step of contacting is performed under conditions and for a time sufficient to permit the removal of cells expressing the protein from the sample.

Within related aspects, methods are provided for inhibiting the development of a cancer in a patient, comprising administering to a patient a biological sample treated as described above.

Methods are further provided, within other aspects, for stimulating and/or expanding T cells specific for a polypeptide of the present invention, comprising contacting T cells with one or more of: (i) a polypeptide as described above; (ii) a polynucleotide encoding such a polypeptide; and/or (iii) an antigen presenting cell that expresses such a polypeptide; under conditions and for a time sufficient to permit the stimulation and/or expansion of T cells. Isolated T cell populations comprising T cells prepared as described above are also provided.

Within further aspects, the present invention provides methods for inhibiting the development of a cancer in a patient, comprising administering to a patient an effective amount of a T cell population as described above.

The present invention further provides methods for inhibiting the development of a cancer in a patient, comprising the steps of: (a) incubating $CD4^+$ and/or $CD8^+$ T cells isolated from a patient with one or more of: (i) a polypeptide comprising at least an immunogenic portion of polypeptide disclosed herein; (ii) a polynucleotide encoding such a polypeptide; and (iii) an antigen-presenting cell that expressed such a polypeptide; and (b) administering to the patient an effective amount of the proliferated T cells, and thereby inhibiting the development of a cancer in the patient. Proliferated cells may, but need not, be cloned prior to administration to the patient.

Within further aspects, the present invention provides methods for determining the presence or absence of a cancer, preferably a colon cancer, in a patient comprising: (a) contacting a biological sample obtained from a patient with a binding agent that binds to a polypeptide as recited above; (b) detecting in the sample an amount of polypeptide that binds to the binding agent; and (c) comparing the amount of polypeptide with a predetermined cut-off value, and therefrom determining the presence or absence of a cancer in the patient. Within preferred embodiments, the binding agent is an antibody, more preferably a monoclonal antibody.

The present invention also provides, within other aspects, methods for monitoring the progression of a cancer in a patient. Such methods comprise the steps of: (a) contacting a biological sample obtained from a patient at a first point in time with a binding agent that binds to a polypeptide as recited above; (b) detecting in the sample an amount of polypeptide that binds to the binding agent; (c) repeating steps (a) and (b) using a biological sample obtained from the patient at a subsequent point in time; and (d) comparing the amount of polypeptide detected in step (c) with the amount detected in step (b) and therefrom monitoring the progression of the cancer in the patient.

The present invention further provides, within other aspects, methods for determining the presence or absence of a cancer in a patient, comprising the steps of: (a)

contacting a biological sample, e.g., tumor sample, serum sample, etc., obtained from a patient with an oligonucleotide that hybridizes to a polynucleotide that encodes a polypeptide of the present invention; (b) detecting in the sample a level of a polynucleotide, preferably mRNA, that hybridizes to the oligonucleotide; and (c) comparing the level of polynucleotide that hybridizes to the oligonucleotide with a predetermined cut-off value, and therefrom determining the presence or absence of a cancer in the patient. Within certain embodiments, the amount of mRNA is detected via polymerase chain reaction using, for example, at least one oligonucleotide primer that hybridizes to a polynucleotide encoding a polypeptide as recited above, or a complement of such a polynucleotide. Within other embodiments, the amount of mRNA is detected using a hybridization technique, employing an oligonucleotide probe that hybridizes to a polynucleotide that encodes a polypeptide as recited above, or a complement of such a polynucleotide.

In related aspects, methods are provided for monitoring the progression of a cancer in a patient, comprising the steps of: (a) contacting a biological sample obtained from a patient with an oligonucleotide that hybridizes to a polynucleotide that encodes a polypeptide of the present invention; (b) detecting in the sample an amount of a polynucleotide that hybridizes to the oligonucleotide; (c) repeating steps (a) and (b) using a biological sample obtained from the patient at a subsequent point in time; and (d) comparing the amount of polynucleotide detected in step (c) with the amount detected in step (b) and therefrom monitoring the progression of the cancer in the patient.

Within further aspects, the present invention provides antibodies, such as monoclonal antibodies, that bind to a polypeptide as described above, as well as diagnostic kits comprising such antibodies. Diagnostic kits comprising one or more oligonucleotide probes or primers as described above are also provided.

These and other aspects of the present invention will become apparent upon reference to the following detailed description. All references disclosed herein are hereby incorporated by reference in their entirety as if each was incorporated individually.

BRIEF DESCRIPTION OF THE SEQUENCE IDENTIFIERS

SEQ ID NO:1 is the determined cDNA sequence for clone R0089:A03

SEQ ID NO:2 is the determined cDNA sequence for clone R0089:A05

SEQ ID NO:3 is the determined cDNA sequence for clone R0089:A06

SEQ ID NO:4 is the determined cDNA sequence for clone R0089:A07

SEQ ID NO:5 is the determined cDNA sequence for clone R0089:A08

SEQ ID NO:6 is the determined cDNA sequence for clone R0089:A09

SEQ ID NO:7 is the determined cDNA sequence for clone R0089:A11

SEQ ID NO:8 is the determined cDNA sequence for clone R0089:A12

SEQ ID NO:9 is the determined cDNA sequence for clone R0089:B02

SEQ ID NO:10 is the determined cDNA sequence for clone R0089:B03

SEQ ID NO:11 is the determined cDNA sequence for clone R0089:B05

SEQ ID NO:12 is the determined cDNA sequence for clone R0089:B06

SEQ ID NO:13 is the determined cDNA sequence for clone R0089:B07

SEQ ID NO:14 is the determined cDNA sequence for clone R0089:B08

SEQ ID NO:15 is the determined cDNA sequence for clone R0089:B09

SEQ ID NO:16 is the determined cDNA sequence for clone R0089:B10

SEQ ID NO:17 is the determined cDNA sequence for clone R0089:B11

SEQ ID NO:18 is the determined cDNA sequence for clone R0089:B12

SEQ ID NO:19 is the determined cDNA sequence for clone R0089:C01

SEQ ID NO:20 is the determined cDNA sequence for clone R0089:C03

SEQ ID NO:21 is the determined cDNA sequence for clone R0089:C04

SEQ ID NO:22 is the determined cDNA sequence for clone R0089:C05

SEQ ID NO:23 is the determined cDNA sequence for clone R0089:C06

SEQ ID NO:24 is the determined cDNA sequence for clone R0089:C07

SEQ ID NO:25 is the determined cDNA sequence for clone R0089:C08

SEQ ID NO:26 is the determined cDNA sequence for clone R0089:C09

SEQ ID NO:27 is the determined cDNA sequence for clone R0089:C10

SEQ ID NO:28 is the determined cDNA sequence for clone R0089:C11

SEQ ID NO:29 is the determined cDNA sequence for clone R0089:C12

SEQ ID NO:30 is the determined cDNA sequence for clone R0089:D01

SEQ ID NO:31 is the determined cDNA sequence for clone R0089:D02
SEQ ID NO:32 is the determined cDNA sequence for clone R0089:D03
SEQ ID NO:33 is the determined cDNA sequence for clone R0089:D04
SEQ ID NO:34 is the determined cDNA sequence for clone R0089:D05
SEQ ID NO:35 is the determined cDNA sequence for clone R0089:D06
SEQ ID NO:36 is the determined cDNA sequence for clone R0089:D07
SEQ ID NO:37 is the determined cDNA sequence for clone R0089:D08
SEQ ID NO:38 is the determined cDNA sequence for clone R0089:D09
SEQ ID NO:39 is the determined cDNA sequence for clone R0089:D10
SEQ ID NO:40 is the determined cDNA sequence for clone R0089:D11
SEQ ID NO:41 is the determined cDNA sequence for clone R0089:D12
SEQ ID NO:42 is the determined cDNA sequence for clone R0089:E01
SEQ ID NO:43 is the determined cDNA sequence for clone R0089:E02
SEQ ID NO:44 is the determined cDNA sequence for clone R0089:E03
SEQ ID NO:45 is the determined cDNA sequence for clone R0089:E05
SEQ ID NO:46 is the determined cDNA sequence for clone R0089:E06
SEQ ID NO:47 is the determined cDNA sequence for clone R0089:E08
SEQ ID NO:48 is the determined cDNA sequence for clone R0089:E09
SEQ ID NO:49 is the determined cDNA sequence for clone R0089:E10
SEQ ID NO:50 is the determined cDNA sequence for clone R0089:E11
SEQ ID NO:51 is the determined cDNA sequence for clone R0089:F01
SEQ ID NO:52 is the determined cDNA sequence for clone R0089:F02
SEQ ID NO:53 is the determined cDNA sequence for clone R0089:F03
SEQ ID NO:54 is the determined cDNA sequence for clone R0089:F04
SEQ ID NO:55 is the determined cDNA sequence for clone R0089:F05
SEQ ID NO:56 is the determined cDNA sequence for clone R0089:F06
SEQ ID NO:57 is the determined cDNA sequence for clone R0089:F07
SEQ ID NO:58 is the determined cDNA sequence for clone R0089:F08
SEQ ID NO:59 is the determined cDNA sequence for clone R0089:F10
SEQ ID NO:60 is the determined cDNA sequence for clone R0089:F11
SEQ ID NO:61 is the determined cDNA sequence for clone R0089:F12

SEQ ID NO:62 is the determined cDNA sequence for clone R0089:G01
SEQ ID NO:63 is the determined cDNA sequence for clone R0089:G02
SEQ ID NO:64 is the determined cDNA sequence for clone R0089:G03
SEQ ID NO:65 is the determined cDNA sequence for clone R0089:G04
SEQ ID NO:66 is the determined cDNA sequence for clone R0089:G05
SEQ ID NO:67 is the determined cDNA sequence for clone R0089:G06
SEQ ID NO:68 is the determined cDNA sequence for clone R0089:G07
SEQ ID NO:69 is the determined cDNA sequence for clone R0089:G09
SEQ ID NO:70 is the determined cDNA sequence for clone R0089:G10
SEQ ID NO:71 is the determined cDNA sequence for clone R0089:G11
SEQ ID NO:72 is the determined cDNA sequence for clone R0089:G12
SEQ ID NO:73 is the determined cDNA sequence for clone R0089:H01
SEQ ID NO:74 is the determined cDNA sequence for clone R0089:H02
SEQ ID NO:75 is the determined cDNA sequence for clone R0089:H03
SEQ ID NO:76 is the determined cDNA sequence for clone R0089:H04
SEQ ID NO:77 is the determined cDNA sequence for clone R0089:H06
SEQ ID NO:78 is the determined cDNA sequence for clone R0089:H07
SEQ ID NO:79 is the determined cDNA sequence for clone R0089:H08
SEQ ID NO:80 is the determined cDNA sequence for clone R0089:H09
SEQ ID NO:81 is the determined cDNA sequence for clone R0089:H10
SEQ ID NO:82 is the determined cDNA sequence for clone R0089:H11
SEQ ID NO:83 is the determined cDNA sequence for clone R0090:A02
SEQ ID NO:84 is the determined cDNA sequence for clone R0090:A05
SEQ ID NO:85 is the determined cDNA sequence for clone R0090:A06
SEQ ID NO:86 is the determined cDNA sequence for clone R0090:A07
SEQ ID NO:87 is the determined cDNA sequence for clone R0090:A08
SEQ ID NO:88 is the determined cDNA sequence for clone R0090:A09
SEQ ID NO:89 is the determined cDNA sequence for clone R0090:A11
SEQ ID NO:90 is the determined cDNA sequence for clone R0090:A12
SEQ ID NO:91 is the determined cDNA sequence for clone R0090:B02
SEQ ID NO:92 is the determined cDNA sequence for clone R0090:B04

SEQ ID NO:93 is the determined cDNA sequence for clone R0090:B05
SEQ ID NO:94 is the determined cDNA sequence for clone R0090:B06
SEQ ID NO:95 is the determined cDNA sequence for clone R0090:B07
SEQ ID NO:96 is the determined cDNA sequence for clone R0090:B08
SEQ ID NO:97 is the determined cDNA sequence for clone R0090:B10
SEQ ID NO:98 is the determined cDNA sequence for clone R0090:B12
SEQ ID NO:99 is the determined cDNA sequence for clone R0090:C01
SEQ ID NO:100 is the determined cDNA sequence for clone R0090:C02
SEQ ID NO:101 is the determined cDNA sequence for clone R0090:C03
SEQ ID NO:102 is the determined cDNA sequence for clone R0090:C04
SEQ ID NO:103 is the determined cDNA sequence for clone R0090:C06
SEQ ID NO:104 is the determined cDNA sequence for clone R0090:C07
SEQ ID NO:105 is the determined cDNA sequence for clone R0090:C08
SEQ ID NO:106 is the determined cDNA sequence for clone R0090:C09
SEQ ID NO:107 is the determined cDNA sequence for clone R0090:C10
SEQ ID NO:108 is the determined cDNA sequence for clone R0090:C11
SEQ ID NO:109 is the determined cDNA sequence for clone R0090:C12
SEQ ID NO:110 is the determined cDNA sequence for clone R0090:D01
SEQ ID NO:111 is the determined cDNA sequence for clone R0090:D02
SEQ ID NO:112 is the determined cDNA sequence for clone R0090:D03
SEQ ID NO:113 is the determined cDNA sequence for clone R0090:D04
SEQ ID NO:114 is the determined cDNA sequence for clone R0090:D05
SEQ ID NO:115 is the determined cDNA sequence for clone R0090:D06
SEQ ID NO:116 is the determined cDNA sequence for clone R0090:D07
SEQ ID NO:117 is the determined cDNA sequence for clone R0090:D08
SEQ ID NO:118 is the determined cDNA sequence for clone R0090:D09
SEQ ID NO:119 is the determined cDNA sequence for clone R0090:D10
SEQ ID NO:120 is the determined cDNA sequence for clone R0090:D11
SEQ ID NO:121 is the determined cDNA sequence for clone R0090:D12
SEQ ID NO:122 is the determined cDNA sequence for clone R0090:E01
SEQ ID NO:123 is the determined cDNA sequence for clone R0090:E02

SEQ ID NO:124 is the determined cDNA sequence for clone R0090:E03
SEQ ID NO:125 is the determined cDNA sequence for clone R0090:E04
SEQ ID NO:126 is the determined cDNA sequence for clone R0090:E05
SEQ ID NO:127 is the determined cDNA sequence for clone R0090:E06
SEQ ID NO:128 is the determined cDNA sequence for clone R0090:E07
SEQ ID NO:129 is the determined cDNA sequence for clone R0090:E08
SEQ ID NO:130 is the determined cDNA sequence for clone R0090:E09
SEQ ID NO:131 is the determined cDNA sequence for clone R0090:E11
SEQ ID NO:132 is the determined cDNA sequence for clone R0090:E12
SEQ ID NO:133 is the determined cDNA sequence for clone R0090:F02
SEQ ID NO:134 is the determined cDNA sequence for clone R0090:F03
SEQ ID NO:135 is the determined cDNA sequence for clone R0090:F04
SEQ ID NO:136 is the determined cDNA sequence for clone R0090:F05
SEQ ID NO:137 is the determined cDNA sequence for clone R0090:F06
SEQ ID NO:138 is the determined cDNA sequence for clone R0090:F07
SEQ ID NO:139 is the determined cDNA sequence for clone R0090:F08
SEQ ID NO:140 is the determined cDNA sequence for clone R0090:F09
SEQ ID NO:141 is the determined cDNA sequence for clone R0090:F10
SEQ ID NO:142 is the determined cDNA sequence for clone R0090:F11
SEQ ID NO:143 is the determined cDNA sequence for clone R0090:F12
SEQ ID NO:144 is the determined cDNA sequence for clone R0090:G01
SEQ ID NO:145 is the determined cDNA sequence for clone R0090:G02
SEQ ID NO:146 is the determined cDNA sequence for clone R0090:G03
SEQ ID NO:147 is the determined cDNA sequence for clone R0090:G05
SEQ ID NO:148 is the determined cDNA sequence for clone R0090:G06
SEQ ID NO:149 is the determined cDNA sequence for clone R0090:G07
SEQ ID NO:150 is the determined cDNA sequence for clone R0090:G08
SEQ ID NO:151 is the determined cDNA sequence for clone R0090:G09
SEQ ID NO:152 is the determined cDNA sequence for clone R0090:G10
SEQ ID NO:153 is the determined cDNA sequence for clone R0090:G11
SEQ ID NO:154 is the determined cDNA sequence for clone R0090:G12

SEQ ID NO:155 is the determined cDNA sequence for clone R0090:H05
SEQ ID NO:156 is the determined cDNA sequence for clone R0090:H06
SEQ ID NO:157 is the determined cDNA sequence for clone R0090:H07
SEQ ID NO:158 is the determined cDNA sequence for clone R0090:H08
SEQ ID NO:159 is the determined cDNA sequence for clone R0090:H09
SEQ ID NO:160 is the determined cDNA sequence for clone R0090:H10
SEQ ID NO:161 is the determined cDNA sequence for clone R0090:H11
SEQ ID NO:162 is the determined cDNA sequence for clone R0091:A02
SEQ ID NO:163 is the determined cDNA sequence for clone R0091:A03
SEQ ID NO:164 is the determined cDNA sequence for clone R0091:A05
SEQ ID NO:165 is the determined cDNA sequence for clone R0091:A06
SEQ ID NO:166 is the determined cDNA sequence for clone R0091:A08
SEQ ID NO:167 is the determined cDNA sequence for clone R0091:A09
SEQ ID NO:168 is the determined cDNA sequence for clone R0091:A11
SEQ ID NO:169 is the determined cDNA sequence for clone R0091:A12
SEQ ID NO:170 is the determined cDNA sequence for clone R0091:B01
SEQ ID NO:171 is the determined cDNA sequence for clone R0091:B02
SEQ ID NO:172 is the determined cDNA sequence for clone R0091:B03
SEQ ID NO:173 is the determined cDNA sequence for clone R0091:B04
SEQ ID NO:174 is the determined cDNA sequence for clone R0091:B05
SEQ ID NO:175 is the determined cDNA sequence for clone R0091:B06
SEQ ID NO:176 is the determined cDNA sequence for clone R0091:B07
SEQ ID NO:177 is the determined cDNA sequence for clone R0091:B08
SEQ ID NO:178 is the determined cDNA sequence for clone R0091:B09
SEQ ID NO:179 is the determined cDNA sequence for clone R0091:B11
SEQ ID NO:180 is the determined cDNA sequence for clone R0091:B12
SEQ ID NO:181 is the determined cDNA sequence for clone R0091:C01
SEQ ID NO:182 is the determined cDNA sequence for clone R0091:C02
SEQ ID NO:183 is the determined cDNA sequence for clone R0091:C03
SEQ ID NO:184 is the determined cDNA sequence for clone R0091:C04
SEQ ID NO:185 is the determined cDNA sequence for clone R0091:C05

SEQ ID NO:186 is the determined cDNA sequence for clone R0091:C06
SEQ ID NO:187 is the determined cDNA sequence for clone R0091:C07
SEQ ID NO:188 is the determined cDNA sequence for clone R0091:C08
SEQ ID NO:189 is the determined cDNA sequence for clone R0091:C09
SEQ ID NO:190 is the determined cDNA sequence for clone R0091:C10
SEQ ID NO:191 is the determined cDNA sequence for clone R0091:C11
SEQ ID NO:192 is the determined cDNA sequence for clone R0091:C12
SEQ ID NO:193 is the determined cDNA sequence for clone R0091:D01
SEQ ID NO:194 is the determined cDNA sequence for clone R0091:D02
SEQ ID NO:195 is the determined cDNA sequence for clone R0091:D03
SEQ ID NO:196 is the determined cDNA sequence for clone R0091:D04
SEQ ID NO:197 is the determined cDNA sequence for clone R0091:D05
SEQ ID NO:198 is the determined cDNA sequence for clone R0091:D06
SEQ ID NO:199 is the determined cDNA sequence for clone R0091:D07
SEQ ID NO:200 is the determined cDNA sequence for clone R0091:D08
SEQ ID NO:201 is the determined cDNA sequence for clone R0091:D09
SEQ ID NO:202 is the determined cDNA sequence for clone R0091:D10
SEQ ID NO:203 is the determined cDNA sequence for clone R0091:D11
SEQ ID NO:204 is the determined cDNA sequence for clone R0091:D12
SEQ ID NO:205 is the determined cDNA sequence for clone R0091:E01
SEQ ID NO:206 is the determined cDNA sequence for clone R0091:E02
SEQ ID NO:207 is the determined cDNA sequence for clone R0091:E03
SEQ ID NO:208 is the determined cDNA sequence for clone R0091:E04
SEQ ID NO:209 is the determined cDNA sequence for clone R0091:E05
SEQ ID NO:210 is the determined cDNA sequence for clone R0091:E06
SEQ ID NO:211 is the determined cDNA sequence for clone R0091:E07
SEQ ID NO:212 is the determined cDNA sequence for clone R0091:E08
SEQ ID NO:213 is the determined cDNA sequence for clone R0091:E09
SEQ ID NO:214 is the determined cDNA sequence for clone R0091:E10
SEQ ID NO:215 is the determined cDNA sequence for clone R0091:E11
SEQ ID NO:216 is the determined cDNA sequence for clone R0091:E12

SEQ ID NO:217 is the determined cDNA sequence for clone R0091:F02
SEQ ID NO:218 is the determined cDNA sequence for clone R0091:F03
SEQ ID NO:219 is the determined cDNA sequence for clone R0091:F04
SEQ ID NO:220 is the determined cDNA sequence for clone R0091:F05
SEQ ID NO:221 is the determined cDNA sequence for clone R0091:F07
SEQ ID NO:222 is the determined cDNA sequence for clone R0091:F09
SEQ ID NO:223 is the determined cDNA sequence for clone R0091:F10
SEQ ID NO:224 is the determined cDNA sequence for clone R0091:F11
SEQ ID NO:225 is the determined cDNA sequence for clone R0091:F12
SEQ ID NO:226 is the determined cDNA sequence for clone R0091:G01
SEQ ID NO:227 is the determined cDNA sequence for clone R0091:G02
SEQ ID NO:228 is the determined cDNA sequence for clone R0091:G04
SEQ ID NO:229 is the determined cDNA sequence for clone R0091:G05
SEQ ID NO:230 is the determined cDNA sequence for clone R0091:G06
SEQ ID NO:231 is the determined cDNA sequence for clone R0091:G07
SEQ ID NO:232 is the determined cDNA sequence for clone R0091:G08
SEQ ID NO:233 is the determined cDNA sequence for clone R0091:G09
SEQ ID NO:234 is the determined cDNA sequence for clone R0091:G10
SEQ ID NO:235 is the determined cDNA sequence for clone R0091:G11
SEQ ID NO:236 is the determined cDNA sequence for clone R0091:G12
SEQ ID NO:237 is the determined cDNA sequence for clone R0091:H01
SEQ ID NO:238 is the determined cDNA sequence for clone R0091:H02
SEQ ID NO:239 is the determined cDNA sequence for clone R0091:H03
SEQ ID NO:240 is the determined cDNA sequence for clone R0091:H04
SEQ ID NO:241 is the determined cDNA sequence for clone R0091:H05
SEQ ID NO:242 is the determined cDNA sequence for clone R0091:H06
SEQ ID NO:243 is the determined cDNA sequence for clone R0091:H07
SEQ ID NO:244 is the determined cDNA sequence for clone R0091:H08
SEQ ID NO:245 is the determined cDNA sequence for clone R0091:H09
SEQ ID NO:246 is the determined cDNA sequence for clone R0091:H10
SEQ ID NO:247 is the determined cDNA sequence for clone R0091:H11

SEQ ID NO:248 is the determined cDNA sequence for clone R0092:A03
SEQ ID NO:249 is the determined cDNA sequence for clone R0092:A05
SEQ ID NO:250 is the determined cDNA sequence for clone R0092:A06
SEQ ID NO:251 is the determined cDNA sequence for clone R0092:A07
SEQ ID NO:252 is the determined cDNA sequence for clone R0092:A09
SEQ ID NO:253 is the determined cDNA sequence for clone R0092:A10
SEQ ID NO:254 is the determined cDNA sequence for clone R0092:A11
SEQ ID NO:255 is the determined cDNA sequence for clone R0092:B01
SEQ ID NO:256 is the determined cDNA sequence for clone R0092:B02
SEQ ID NO:257 is the determined cDNA sequence for clone R0092:B03
SEQ ID NO:258 is the determined cDNA sequence for clone R0092:B04
SEQ ID NO:259 is the determined cDNA sequence for clone R0092:B05
SEQ ID NO:260 is the determined cDNA sequence for clone R0092:B08
SEQ ID NO:261 is the determined cDNA sequence for clone R0092:B09
SEQ ID NO:262 is the determined cDNA sequence for clone R0092:B10
SEQ ID NO:263 is the determined cDNA sequence for clone R0092:B11
SEQ ID NO:264 is the determined cDNA sequence for clone R0092:B12
SEQ ID NO:265 is the determined cDNA sequence for clone R0092:C02
SEQ ID NO:266 is the determined cDNA sequence for clone R0092:C03
SEQ ID NO:267 is the determined cDNA sequence for clone R0092:C04
SEQ ID NO:268 is the determined cDNA sequence for clone R0092:C05
SEQ ID NO:269 is the determined cDNA sequence for clone R0092:C06
SEQ ID NO:270 is the determined cDNA sequence for clone R0092:C07
SEQ ID NO:271 is the determined cDNA sequence for clone R0092:C08
SEQ ID NO:272 is the determined cDNA sequence for clone R0092:C09
SEQ ID NO:273 is the determined cDNA sequence for clone R0092:C10
SEQ ID NO:274 is the determined cDNA sequence for clone R0092:C11
SEQ ID NO:275 is the determined cDNA sequence for clone R0092:C12
SEQ ID NO:276 is the determined cDNA sequence for clone R0092:D02
SEQ ID NO:277 is the determined cDNA sequence for clone R0092:D03
SEQ ID NO:278 is the determined cDNA sequence for clone R0092:C04

SEQ ID NO:279 is the determined cDNA sequence for clone R0092:D05
SEQ ID NO:280 is the determined cDNA sequence for clone R0092:D06
SEQ ID NO:281 is the determined cDNA sequence for clone R0092:D07
SEQ ID NO:282 is the determined cDNA sequence for clone R0092:D08
SEQ ID NO:283 is the determined cDNA sequence for clone R0092:D09
SEQ ID NO:284 is the determined cDNA sequence for clone R0092:D10
SEQ ID NO:285 is the determined cDNA sequence for clone R0092:D11
SEQ ID NO:286 is the determined cDNA sequence for clone R0092:D12
SEQ ID NO:287 is the determined cDNA sequence for clone R0092:E01
SEQ ID NO:288 is the determined cDNA sequence for clone R0092:E02
SEQ ID NO:289 is the determined cDNA sequence for clone R0092:E03
SEQ ID NO:290 is the determined cDNA sequence for clone R0092:E04
SEQ ID NO:291 is the determined cDNA sequence for clone R0092:E05
SEQ ID NO:292 is the determined cDNA sequence for clone R0092:E06
SEQ ID NO:293 is the determined cDNA sequence for clone R0092:E07
SEQ ID NO:294 is the determined cDNA sequence for clone R0092:E08
SEQ ID NO:295 is the determined cDNA sequence for clone R0092:E09
SEQ ID NO:296 is the determined cDNA sequence for clone R0092:E10
SEQ ID NO:297 is the determined cDNA sequence for clone R0092:E11
SEQ ID NO:298 is the determined cDNA sequence for clone R0092:E12
SEQ ID NO:299 is the determined cDNA sequence for clone R0092:F01
SEQ ID NO:300 is the determined cDNA sequence for clone R0092:F02
SEQ ID NO:301 is the determined cDNA sequence for clone R0092:F03
SEQ ID NO:302 is the determined cDNA sequence for clone R0092:F04
SEQ ID NO:303 is the determined cDNA sequence for clone R0092:F05
SEQ ID NO:304 is the determined cDNA sequence for clone R0092:F06
SEQ ID NO:305 is the determined cDNA sequence for clone R0092:F07
SEQ ID NO:306 is the determined cDNA sequence for clone R0092:F08
SEQ ID NO:307 is the determined cDNA sequence for clone R0092:F09
SEQ ID NO:308 is the determined cDNA sequence for clone R0092:F10
SEQ ID NO:309 is the determined cDNA sequence for clone R0092:F11

SEQ ID NO:310 is the determined cDNA sequence for clone R0092:F12

SEQ ID NO:311 is the determined cDNA sequence for clone R0092:G01

SEQ ID NO:312 is the determined cDNA sequence for clone R0092:G02

SEQ ID NO:313 is the determined cDNA sequence for clone R0092:G03

SEQ ID NO:314 is the determined cDNA sequence for clone R0092:G04

SEQ ID NO:315 is the determined cDNA sequence for clone R0092:G05

SEQ ID NO:316 is the determined cDNA sequence for clone R0092:G06

SEQ ID NO:317 is the determined cDNA sequence for clone R0092:G07

SEQ ID NO:318 is the determined cDNA sequence for clone R0092:G08

SEQ ID NO:319 is the determined cDNA sequence for clone R0092:G09

SEQ ID NO:320 is the determined cDNA sequence for clone R0092:G10

SEQ ID NO:321 is the determined cDNA sequence for clone R0092:G11

SEQ ID NO:322 is the determined cDNA sequence for clone R0092:G12

SEQ ID NO:323 is the determined cDNA sequence for clone R0092:H01

SEQ ID NO:324 is the determined cDNA sequence for clone R0092:H02

SEQ ID NO:325 is the determined cDNA sequence for clone R0092:H03

SEQ ID NO:326 is the determined cDNA sequence for clone R0092:H04

SEQ ID NO:327 is the determined cDNA sequence for clone R0092:H05

SEQ ID NO:328 is the determined cDNA sequence for clone R0092:H06

SEQ ID NO:329 is the determined cDNA sequence for clone R0092:H07

SEQ ID NO:330 is the determined cDNA sequence for clone R0092:H08

SEQ ID NO:331 is the determined cDNA sequence for clone R0092:H09

SEQ ID NO:332 is the determined cDNA sequence for clone R0092:H10

SEQ ID NO:333 is the determined cDNA sequence for clone R0092:H11

SEQ ID NO:334 is the determined cDNA sequence for a clone from a primary normal colon library

SEQ ID NO:335 is the determined cDNA sequence for clone 89A9_C1410P

SEQ ID NO:336 is the determined cDNA sequence for clone 89C4_C1411P

SEQ ID NO:337 is the determined cDNA sequence for clone 89E2_C1412P

SEQ ID NO:338 is the determined cDNA sequence for clone 89G10_C1413P

SEQ ID NO:339 is the determined cDNA sequence for clone 89G2_C1407P

SEQ ID NO:340 is the determined cDNA sequence for clone 90C11_C1414P

SEQ ID NO:341 is the determined cDNA sequence for clone 90F8_C1408P

SEQ ID NO:342 is the determined cDNA sequence for clone 90H10_C1415P

SEQ ID NO:343 is the determined cDNA sequence for clone 91D6_C1416P

SEQ ID NO:344 is the determined cDNA sequence for clone 92B4_C1409P

SEQ ID NO:345 is the determined cDNA sequence for clone 92H6_C1417P

SEQ ID NO:346 is the determined cDNA sequence for clone 93F10_C1418P

SEQ ID NO:347 is the determined cDNA sequence for clone 94E8_C1419P

SEQ ID NO:348 is the determined cDNA sequence for clone 95D1_c592S

SEQ ID NO:349 is the determined cDNA sequence for clone 98F12_C1421P

SEQ ID NO:350 is the determined cDNA sequence for clone 98H6

SEQ ID NO:351 is the determined cDNA sequence for clone 99E5_C1401P

SEQ ID NO:352 is the determined cDNA sequence for clone 100G8_C1422P

SEQ ID NO:353 is the determined cDNA sequence for clone 101G6_C1402P

SEQ ID NO:354 is the determined cDNA sequence for clone 103F6

SEQ ID NO:355 is the determined cDNA sequence for clone 104C9_C1404P

SEQ ID NO:356 is the determined cDNA sequence for clone 109C2_C1405P

SEQ ID NO:357 is the determined cDNA sequence for clone 109E8_C1406P

SEQ ID NO:358 is the determined cDNA sequence for clone 95A4

SEQ ID NO:359 is the determined cDNA sequence for clone 93F12

SEQ ID NO:360 is the determined cDNA sequence for clone 93H11

SEQ ID NO:361 is the determined cDNA sequence for clone 110D9

SEQ ID NO:362 is the determined cDNA sequence for clone 102E7

SEQ ID NO:363 is the determined cDNA sequence for clone '59698.1

SEQ ID NO:364 is the determined cDNA sequence for clone '59699.3

SEQ ID NO:365 is the determined cDNA sequence for clone '59717.2

SEQ ID NO:366 is the determined cDNA sequence for clone '59717.4

SEQ ID NO:367 is the determined cDNA sequence for clone '59719.2

SEQ ID NO:368 is the determined cDNA sequence for clone '59719.4

SEQ ID NO:369 is the determined cDNA sequence for clone '59720.1

SEQ ID NO:370 is the determined cDNA sequence for clone '59721.1

SEQ ID NO:371 is the determined cDNA sequence for clone '60768.1

SEQ ID NO:372 is the determined cDNA sequence for clone '60769.1

SEQ ID NO:373 is the determined cDNA sequence for clone '60770.1

SEQ ID NO:374 is the determined cDNA sequence for clone '60773.1

SEQ ID NO:375 is the determined cDNA sequence for clone '60776.1

SEQ ID NO:376 is the determined cDNA sequence for clone '60777.1

SEQ ID NO:377 is the determined cDNA sequence for clone '60778.1

DETAILED DESCRIPTION OF THE INVENTION

The present invention is directed generally to compositions and their use in the therapy and diagnosis of cancer, particularly colon cancer. As described further below, illustrative compositions of the present invention include, but are not restricted to, polypeptides, particularly immunogenic polypeptides, polynucleotides encoding such polypeptides, antibodies and other binding agents, antigen presenting cells (APCs) and immune system cells (*e.g.*, T cells).

The practice of the present invention will employ, unless indicated specifically to the contrary, conventional methods of virology, immunology, microbiology, molecular biology and recombinant DNA techniques within the skill of the art, many of which are described below for the purpose of illustration. Such techniques are explained fully in the literature. See, *e.g.*, Sambrook, et al. Molecular Cloning: A Laboratory Manual (2nd Edition, 1989); Maniatis et al. Molecular Cloning: A Laboratory Manual (1982); DNA Cloning: A Practical Approach, vol. I & II (D. Glover, ed.); Oligonucleotide Synthesis (N. Gait, ed., 1984); Nucleic Acid Hybridization (B. Hames & S. Higgins, eds., 1985); Transcription and Translation (B. Hames & S. Higgins, eds., 1984); Animal Cell Culture (R. Freshney, ed., 1986); Perbal, A Practical Guide to Molecular Cloning (1984).

All publications, patents and patent applications cited herein, whether supra or infra, are hereby incorporated by reference in their entirety.

As used in this specification and the appended claims, the singular forms "a," "an" and "the" include plural references unless the content clearly dictates otherwise.

POLYPEPTIDE COMPOSITIONS

As used herein, the term "polypeptide" " is used in its conventional meaning, *i.e.*, as a sequence of amino acids. The polypeptides are not limited to a specific length of the product; thus, peptides, oligopeptides, and proteins are included within the definition of polypeptide, and such terms may be used interchangeably herein unless specifically indicated otherwise. This term also does not refer to or exclude post-expression modifications of the polypeptide, for example, glycosylations,

acetylations, phosphorylations and the like, as well as other modifications known in the art, both naturally occurring and non-naturally occurring. A polypeptide may be an entire protein, or a subsequence thereof. Particular polypeptides of interest in the context of this invention are amino acid subsequences comprising epitopes, *i*.e., antigenic determinants substantially responsible for the immunogenic properties of a polypeptide and being capable of evoking an immune response.

Particularly illustrative polypeptides of the present invention comprise those encoded by a polynucleotide sequence set forth in any one of SEQ ID NO:1-377, or a sequence that hybridizes under moderately stringent conditions, or, alternatively, under highly stringent conditions, to a polynucleotide sequence set forth in any one of SEQ ID NO:1-377.

The polypeptides of the present invention are sometimes herein referred to as colon tumor proteins or colon tumor polypeptides, as an indication that their identification has been based at least in part upon their increased levels of expression in colon tumor samples. Thus, a "colon tumor polypeptide" or "colon tumor protein," refers generally to a polypeptide sequence of the present invention, or a polynucleotide sequence encoding such a polypeptide, that is expressed in a substantial proportion of colon tumor samples, for example preferably greater than about 20%, more preferably greater than about 30%, and most preferably greater than about 50% or more of colon tumor samples tested, at a level that is at least two fold, and preferably at least five fold, greater than the level of expression in normal tissues, as determined using a representative assay provided herein. A colon tumor polypeptide sequence of the invention, based upon its increased level of expression in tumor cells, has particular utility both as a diagnostic marker as well as a therapeutic target, as further described below.

In certain preferred embodiments, the polypeptides of the invention are immunogenic, *i.e.*, they react detectably within an immunoassay (such as an ELISA or T-cell stimulation assay) with antisera and/or T-cells from a patient with colon cancer. Screening for immunogenic activity can be performed using techniques well known to the skilled artisan. For example, such screens can be performed using methods such as those described in Harlow and Lane, *Antibodies: A Laboratory Manual,* Cold Spring

Harbor Laboratory, 1988. In one illustrative example, a polypeptide may be immobilized on a solid support and contacted with patient sera to allow binding of antibodies within the sera to the immobilized polypeptide. Unbound sera may then be removed and bound antibodies detected using, for example, ^{125}I-labeled Protein A.

As would be recognized by the skilled artisan, immunogenic portions of the polypeptides disclosed herein are also encompassed by the present invention. An "immunogenic portion," as used herein, is a fragment of an immunogenic polypeptide of the invention that itself is immunologically reactive (*i.e.,* specifically binds) with the B-cells and/or T-cell surface antigen receptors that recognize the polypeptide. Immunogenic portions may generally be identified using well known techniques, such as those summarized in Paul, *Fundamental Immunology*, 3rd ed., 243-247 (Raven Press, 1993) and references cited therein. Such techniques include screening polypeptides for the ability to react with antigen-specific antibodies, antisera and/or T-cell lines or clones. As used herein, antisera and antibodies are "antigen-specific" if they specifically bind to an antigen (*i.e.,* they react with the protein in an ELISA or other immunoassay, and do not react detectably with unrelated proteins). Such antisera and antibodies may be prepared as described herein, and using well-known techniques.

In one preferred embodiment, an immunogenic portion of a polypeptide of the present invention is a portion that reacts with antisera and/or T-cells at a level that is not substantially less than the reactivity of the full-length polypeptide (*e.g.,* in an ELISA and/or T-cell reactivity assay). Preferably, the level of immunogenic activity of the immunogenic portion is at least about 50%, preferably at least about 70% and most preferably greater than about 90% of the immunogenicity for the full-length polypeptide. In some instances, preferred immunogenic portions will be identified that have a level of immunogenic activity greater than that of the corresponding full-length polypeptide, *e.g.*, having greater than about 100% or 150% or more immunogenic activity.

In certain other embodiments, illustrative immunogenic portions may include peptides in which an N-terminal leader sequence and/or transmembrane domain have been deleted. Other illustrative immunogenic portions will contain a small N-

and/or C-terminal deletion (*e.g.*, 1-30 amino acids, preferably 5-15 amino acids), relative to the mature protein.

In another embodiment, a polypeptide composition of the invention may also comprise one or more polypeptides that are immunologically reactive with T cells and/or antibodies generated against a polypeptide of the invention, particularly a polypeptide having an amino acid sequence disclosed herein, or to an immunogenic fragment or variant thereof.

In another embodiment of the invention, polypeptides are provided that comprise one or more polypeptides that are capable of eliciting T cells and/or antibodies that are immunologically reactive with one or more polypeptides described herein, or one or more polypeptides encoded by contiguous nucleic acid sequences contained in the polynucleotide sequences disclosed herein, or immunogenic fragments or variants thereof, or to one or more nucleic acid sequences which hybridize to one or more of these sequences under conditions of moderate to high stringency.

The present invention, in another aspect, provides polypeptide fragments comprising at least about 5, 10, 15, 20, 25, 50, or 100 contiguous amino acids, or more, including all intermediate lengths, of a polypeptide compositions set forth herein, such as those encoded by a polynucleotide sequence set forth in a sequence of SEQ ID NO:1-377.

In another aspect, the present invention provides variants of the polypeptide compositions described herein. Polypeptide variants generally encompassed by the present invention will typically exhibit at least about 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% or more identity (determined as described below), along its length, to a polypeptide sequences set forth herein.

In one preferred embodiment, the polypeptide fragments and variants provided by the present invention are immunologically reactive with an antibody and/or T-cell that reacts with a full-length polypeptide specifically set forth herein.

In another preferred embodiment, the polypeptide fragments and variants provided by the present invention exhibit a level of immunogenic activity of at least about 50%, preferably at least about 70%, and most preferably at least about 90% or

more of that exhibited by a full-length polypeptide sequence specifically set forth herein.

A polypeptide "variant," as the term is used herein, is a polypeptide that typically differs from a polypeptide specifically disclosed herein in one or more substitutions, deletions, additions and/or insertions. Such variants may be naturally occurring or may be synthetically generated, for example, by modifying one or more of the above polypeptide sequences of the invention and evaluating their immunogenic activity as described herein and/or using any of a number of techniques well known in the art.

For example, certain illustrative variants of the polypeptides of the invention include those in which one or more portions, such as an N-terminal leader sequence or transmembrane domain, have been removed. Other illustrative variants include variants in which a small portion (*e.g.,* 1-30 amino acids, preferably 5-15 amino acids) has been removed from the N- and/or C-terminal of the mature protein.

In many instances, a variant will contain conservative substitutions. A "conservative substitution" is one in which an amino acid is substituted for another amino acid that has similar properties, such that one skilled in the art of peptide chemistry would expect the secondary structure and hydropathic nature of the polypeptide to be substantially unchanged. As described above, modifications may be made in the structure of the polynucleotides and polypeptides of the present invention and still obtain a functional molecule that encodes a variant or derivative polypeptide with desirable characteristics, *e.g.*, with immunogenic characteristics. When it is desired to alter the amino acid sequence of a polypeptide to create an equivalent, or even an improved, immunogenic variant or portion of a polypeptide of the invention, one skilled in the art will typically change one or more of the codons of the encoding DNA sequence according to Table 1.

For example, certain amino acids may be substituted for other amino acids in a protein structure without appreciable loss of interactive binding capacity with structures such as, for example, antigen-binding regions of antibodies or binding sites on substrate molecules. Since it is the interactive capacity and nature of a protein that defines that protein's biological functional activity, certain amino acid sequence

substitutions can be made in a protein sequence, and, of course, its underlying DNA coding sequence, and nevertheless obtain a protein with like properties. It is thus contemplated that various changes may be made in the peptide sequences of the disclosed compositions, or corresponding DNA sequences which encode said peptides without appreciable loss of their biological utility or activity.

TABLE 1

Amino Acids			Codons					
Alanine	Ala	A	GCA	GCC	GCG	GCU		
Cysteine	Cys	C	UGC	UGU				
Aspartic acid	Asp	D	GAC	GAU				
Glutamic acid	Glu	E	GAA	GAG				
Phenylalanine	Phe	F	UUC	UUU				
Glycine	Gly	G	GGA	GGC	GGG	GGU		
Histidine	His	H	CAC	CAU				
Isoleucine	Ile	I	AUA	AUC	AUU			
Lysine	Lys	K	AAA	AAG				
Leucine	Leu	L	UUA	UUG	CUA	CUC	CUG	CUU
Methionine	Met	M	AUG					
Asparagine	Asn	N	AAC	AAU				
Proline	Pro	P	CCA	CCC	CCG	CCU		
Glutamine	Gln	Q	CAA	CAG				
Arginine	Arg	R	AGA	AGG	CGA	CGC	CGG	CGU
Serine	Ser	S	AGC	AGU	UCA	UCC	UCG	UCU
Threonine	Thr	T	ACA	ACC	ACG	ACU		
Valine	Val	V	GUA	GUC	GUG	GUU		
Tryptophan	Trp	W	UGG					
Tyrosine	Tyr	Y	UAC	UAU				

In making such changes, the hydropathic index of amino acids may be considered. The importance of the hydropathic amino acid index in conferring interactive biologic function on a protein is generally understood in the art (Kyte and

Doolittle, 1982, incorporated herein by reference). It is accepted that the relative hydropathic character of the amino acid contributes to the secondary structure of the resultant protein, which in turn defines the interaction of the protein with other molecules, for example, enzymes, substrates, receptors, DNA, antibodies, antigens, and the like. Each amino acid has been assigned a hydropathic index on the basis of its hydrophobicity and charge characteristics (Kyte and Doolittle, 1982). These values are: isoleucine (+4.5); valine (+4.2); leucine (+3.8); phenylalanine (+2.8); cysteine/cystine (+2.5); methionine (+1.9); alanine (+1.8); glycine (–0.4); threonine (–0.7); serine (–0.8); tryptophan (–0.9); tyrosine (–1.3); proline (–1.6); histidine (–3.2); glutamate (–3.5); glutamine (–3.5); aspartate (–3.5); asparagine (–3.5); lysine (–3.9); and arginine (–4.5).

It is known in the art that certain amino acids may be substituted by other amino acids having a similar hydropathic index or score and still result in a protein with similar biological activity, *i.e.* still obtain a biological functionally equivalent protein. In making such changes, the substitution of amino acids whose hydropathic indices are within ±2 is preferred, those within ±1 are particularly preferred, and those within ±0.5 are even more particularly preferred. It is also understood in the art that the substitution of like amino acids can be made effectively on the basis of hydrophilicity. U. S. Patent 4,554,101 (specifically incorporated herein by reference in its entirety), states that the greatest local average hydrophilicity of a protein, as governed by the hydrophilicity of its adjacent amino acids, correlates with a biological property of the protein.

As detailed in U. S. Patent 4,554,101, the following hydrophilicity values have been assigned to amino acid residues: arginine (+3.0); lysine (+3.0); aspartate (+3.0 ± 1); glutamate (+3.0 ± 1); serine (+0.3); asparagine (+0.2); glutamine (+0.2); glycine (0); threonine (–0.4); proline (–0.5 ± 1); alanine (–0.5); histidine (–0.5); cysteine (–1.0); methionine (–1.3); valine (–1.5); leucine (–1.8); isoleucine (–1.8); tyrosine (–2.3); phenylalanine (–2.5); tryptophan (–3.4). It is understood that an amino acid can be substituted for another having a similar hydrophilicity value and still obtain a biologically equivalent, and in particular, an immunologically equivalent protein. In such changes, the substitution of amino acids whose hydrophilicity values are within ±2

is preferred, those within ±1 are particularly preferred, and those within ±0.5 are even more particularly preferred.

As outlined above, amino acid substitutions are generally therefore based on the relative similarity of the amino acid side-chain substituents, for example, their hydrophobicity, hydrophilicity, charge, size, and the like. Exemplary substitutions that take various of the foregoing characteristics into consideration are well known to those of skill in the art and include: arginine and lysine; glutamate and aspartate; serine and threonine; glutamine and asparagine; and valine, leucine and isoleucine.

In addition, any polynucleotide may be further modified to increase stability *in vivo*. Possible modifications include, but are not limited to, the addition of flanking sequences at the 5' and/or 3' ends; the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages in the backbone; and/or the inclusion of nontraditional bases such as inosine, queosine and wybutosine, as well as acetyl-methyl-, thio- and other modified forms of adenine, cytidine, guanine, thymine and uridine.

Amino acid substitutions may further be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity and/or the amphipathic nature of the residues. For example, negatively charged amino acids include aspartic acid and glutamic acid; positively charged amino acids include lysine and arginine; and amino acids with uncharged polar head groups having similar hydrophilicity values include leucine, isoleucine and valine; glycine and alanine; asparagine and glutamine; and serine, threonine, phenylalanine and tyrosine. Other groups of amino acids that may represent conservative changes include: (1) ala, pro, gly, glu, asp, gln, asn, ser, thr; (2) cys, ser, tyr, thr; (3) val, ile, leu, met, ala, phe; (4) lys, arg, his; and (5) phe, tyr, trp, his. A variant may also, or alternatively, contain nonconservative changes. In a preferred embodiment, variant polypeptides differ from a native sequence by substitution, deletion or addition of five amino acids or fewer. Variants may also (or alternatively) be modified by, for example, the deletion or addition of amino acids that have minimal influence on the immunogenicity, secondary structure and hydropathic nature of the polypeptide.

As noted above, polypeptides may comprise a signal (or leader) sequence at the N-terminal end of the protein, which co-translationally or post-translationally directs transfer of the protein. The polypeptide may also be conjugated to a linker or other sequence for ease of synthesis, purification or identification of the polypeptide (*e.g.,* poly-His), or to enhance binding of the polypeptide to a solid support. For example, a polypeptide may be conjugated to an immunoglobulin Fc region.

When comparing polypeptide sequences, two sequences are said to be "identical" if the sequence of amino acids in the two sequences is the same when aligned for maximum correspondence, as described below. Comparisons between two sequences are typically performed by comparing the sequences over a comparison window to identify and compare local regions of sequence similarity. A "comparison window" as used herein, refers to a segment of at least about 20 contiguous positions, usually 30 to about 75, 40 to about 50, in which a sequence may be compared to a reference sequence of the same number of contiguous positions after the two sequences are optimally aligned.

Optimal alignment of sequences for comparison may be conducted using the Megalign program in the Lasergene suite of bioinformatics software (DNASTAR, Inc., Madison, WI), using default parameters. This program embodies several alignment schemes described in the following references: Dayhoff, M.O. (1978) A model of evolutionary change in proteins – Matrices for detecting distant relationships. In Dayhoff, M.O. (ed.) Atlas of Protein Sequence and Structure, National Biomedical Research Foundation, Washington DC Vol. 5, Suppl. 3, pp. 345-358; Hein J. (1990) Unified Approach to Alignment and Phylogenes pp. 626-645 *Methods in Enzymology* vol. 183, Academic Press, Inc., San Diego, CA; Higgins, D.G. and Sharp, P.M. (1989) *CABIOS 5*:151-153; Myers, E.W. and Muller W. (1988) *CABIOS 4*:11-17; Robinson, E.D. (1971) *Comb. Theor 11*:105; Saitou, N. Nei, M. (1987) *Mol. Biol. Evol. 4*:406-425; Sneath, P.H.A. and Sokal, R.R. (1973) *Numerical Taxonomy – the Principles and Practice of Numerical Taxonomy*, Freeman Press, San Francisco, CA; Wilbur, W.J. and Lipman, D.J. (1983) *Proc. Natl. Acad., Sci. USA 80*:726-730.

Alternatively, optimal alignment of sequences for comparison may be conducted by the local identity algorithm of Smith and Waterman (1981) *Add. APL.*

Math 2:482, by the identity alignment algorithm of Needleman and Wunsch (1970) *J. Mol. Biol.* 48:443, by the search for similarity methods of Pearson and Lipman (1988) *Proc. Natl. Acad. Sci. USA* 85: 2444, by computerized implementations of these algorithms (GAP, BESTFIT, BLAST, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group (GCG), 575 Science Dr., Madison, WI), or by inspection.

One preferred example of algorithms that are suitable for determining percent sequence identity and sequence similarity are the BLAST and BLAST 2.0 algorithms, which are described in Altschul et al. (1977) *Nucl. Acids Res.* 25:3389-3402 and Altschul et al. (1990) *J. Mol. Biol.* 215:403-410, respectively. BLAST and BLAST 2.0 can be used, for example with the parameters described herein, to determine percent sequence identity for the polynucleotides and polypeptides of the invention. Software for performing BLAST analyses is publicly available through the National Center for Biotechnology Information. For amino acid sequences, a scoring matrix can be used to calculate the cumulative score. Extension of the word hits in each direction are halted when: the cumulative alignment score falls off by the quantity X from its maximum achieved value; the cumulative score goes to zero or below, due to the accumulation of one or more negative-scoring residue alignments; or the end of either sequence is reached. The BLAST algorithm parameters W, T and X determine the sensitivity and speed of the alignment.

In one preferred approach, the "percentage of sequence identity" is determined by comparing two optimally aligned sequences over a window of comparison of at least 20 positions, wherein the portion of the polypeptide sequence in the comparison window may comprise additions or deletions (*i.e.,* gaps) of 20 percent or less, usually 5 to 15 percent, or 10 to 12 percent, as compared to the reference sequences (which does not comprise additions or deletions) for optimal alignment of the two sequences. The percentage is calculated by determining the number of positions at which the identical amino acid residue occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the reference sequence (*i.e.,* the window size) and multiplying the results by 100 to yield the percentage of sequence identity.

Within other illustrative embodiments, a polypeptide may be a xenogeneic polypeptide that comprises an polypeptide having substantial sequence identity, as described above, to the human polypeptide (also termed autologous antigen) which served as a reference polypeptide, but which xenogeneic polypeptide is derived from a different, non-human species. One skilled in the art will recognize that "self"antigens are often poor stimulators of CD8+ and CD4+ T-lymphocyte responses, and therefore efficient immunotherapeutic strategies directed against tumor polypeptides require the development of methods to overcome immune tolerance to particular self tumor polypeptides. For example, humans immunized with prostase protein from a xenogeneic (non human) origin are capable of mounting an immune response against the counterpart human protein, *e.g.* the human prostase tumor protein present on human tumor cells. Accordingly, the present invention provides methods for purifying the xenogeneic form of the tumor proteins set forth herein, such as the polypeptides encoded by polynucleotide sequences set forth in SEQ ID NO:1-377.

Therefore, one aspect of the present invention provides xenogeneic variants of the polypeptide compositions described herein. Such xenogeneic variants generally encompassed by the present invention will typically exhibit at least about 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% or more identity along their lengths, to a polypeptide sequences set forth herein.

More particularly, the invention is directed to mouse, rat, monkey, porcine and other non-human polypeptides which can be used as xenogeneic forms of human polypeptides set forth herein, to induce immune responses directed against tumor polypeptides of the invention.

Within other illustrative embodiments, a polypeptide may be a fusion polypeptide that comprises multiple polypeptides as described herein, or that comprises at least one polypeptide as described herein and an unrelated sequence, such as a known tumor protein. A fusion partner may, for example, assist in providing T helper epitopes (an immunological fusion partner), preferably T helper epitopes recognized by humans, or may assist in expressing the protein (an expression enhancer) at higher yields than the native recombinant protein. Certain preferred fusion partners are both immunological and expression enhancing fusion partners. Other fusion partners may be

selected so as to increase the solubility of the polypeptide or to enable the polypeptide to be targeted to desired intracellular compartments. Still further fusion partners include affinity tags, which facilitate purification of the polypeptide.

Fusion polypeptides may generally be prepared using standard techniques, including chemical conjugation. Preferably, a fusion polypeptide is expressed as a recombinant polypeptide, allowing the production of increased levels, relative to a non-fused polypeptide, in an expression system. Briefly, DNA sequences encoding the polypeptide components may be assembled separately, and ligated into an appropriate expression vector. The 3' end of the DNA sequence encoding one polypeptide component is ligated, with or without a peptide linker, to the 5' end of a DNA sequence encoding the second polypeptide component so that the reading frames of the sequences are in phase. This permits translation into a single fusion polypeptide that retains the biological activity of both component polypeptides.

A peptide linker sequence may be employed to separate the first and second polypeptide components by a distance sufficient to ensure that each polypeptide folds into its secondary and tertiary structures. Such a peptide linker sequence is incorporated into the fusion polypeptide using standard techniques well known in the art. Suitable peptide linker sequences may be chosen based on the following factors: (1) their ability to adopt a flexible extended conformation; (2) their inability to adopt a secondary structure that could interact with functional epitopes on the first and second polypeptides; and (3) the lack of hydrophobic or charged residues that might react with the polypeptide functional epitopes. Preferred peptide linker sequences contain Gly, Asn and Ser residues. Other near neutral amino acids, such as Thr and Ala may also be used in the linker sequence. Amino acid sequences which may be usefully employed as linkers include those disclosed in Maratea et al., *Gene 40*:39-46, 1985; Murphy et al., *Proc. Natl. Acad. Sci. USA 83*:8258-8262, 1986; U.S. Patent No. 4,935,233 and U.S. Patent No. 4,751,180. The linker sequence may generally be from 1 to about 50 amino acids in length. Linker sequences are not required when the first and second polypeptides have non-essential N-terminal amino acid regions that can be used to separate the functional domains and prevent steric interference.

The ligated DNA sequences are operably linked to suitable transcriptional or translational regulatory elements. The regulatory elements responsible for expression of DNA are located only 5' to the DNA sequence encoding the first polypeptides. Similarly, stop codons required to end translation and transcription termination signals are only present 3' to the DNA sequence encoding the second polypeptide.

The fusion polypeptide can comprise a polypeptide as described herein together with an unrelated immunogenic protein, such as an immunogenic protein capable of eliciting a recall response. Examples of such proteins include tetanus, tuberculosis and hepatitis proteins (*see*, for example, Stoute et al. *New Engl. J. Med.*, *336*:86-91, 1997).

In one preferred embodiment, the immunological fusion partner is derived from a Mycobacterium sp., such as a Mycobacterium tuberculosis-derived Ra12 fragment. Ra12 compositions and methods for their use in enhancing the expression and/or immunogenicity of heterologous polynucleotide/polypeptide sequences is described in U.S. Patent Application 60/158,585, the disclosure of which is incorporated herein by reference in its entirety. Briefly, Ra12 refers to a polynucleotide region that is a subsequence of a *Mycobacterium tuberculosis* MTB32A nucleic acid. MTB32A is a serine protease of 32 KD molecular weight encoded by a gene in virulent and avirulent strains of *M. tuberculosis*. The nucleotide sequence and amino acid sequence of MTB32A have been described (for example, U.S. Patent Application 60/158,585; see also, Skeiky *et al.*, *Infection and Immun.* (1999) 67:3998-4007, incorporated herein by reference). C-terminal fragments of the MTB32A coding sequence express at high levels and remain as a soluble polypeptides throughout the purification process. Moreover, Ra12 may enhance the immunogenicity of heterologous immunogenic polypeptides with which it is fused. One preferred Ra12 fusion polypeptide comprises a 14 KD C-terminal fragment corresponding to amino acid residues 192 to 323 of MTB32A. Other preferred Ra12 polynucleotides generally comprise at least about 15 consecutive nucleotides, at least about 30 nucleotides, at least about 60 nucleotides, at least about 100 nucleotides, at least about 200 nucleotides, or at least about 300 nucleotides that encode a portion of a Ra12 polypeptide. Ra12

polynucleotides may comprise a native sequence (*i.e.,* an endogenous sequence that encodes a Ra12 polypeptide or a portion thereof) or may comprise a variant of such a sequence. Ra12 polynucleotide variants may contain one or more substitutions, additions, deletions and/or insertions such that the biological activity of the encoded fusion polypeptide is not substantially diminished, relative to a fusion polypeptide comprising a native Ra12 polypeptide. Variants preferably exhibit at least about 70% identity, more preferably at least about 80% identity and most preferably at least about 90% identity to a polynucleotide sequence that encodes a native Ra12 polypeptide or a portion thereof.

Within other preferred embodiments, an immunological fusion partner is derived from protein D, a surface protein of the gram-negative bacterium Haemophilus influenza B (WO 91/18926). Preferably, a protein D derivative comprises approximately the first third of the protein (*e.g.,* the first N-terminal 100-110 amino acids), and a protein D derivative may be lipidated. Within certain preferred embodiments, the first 109 residues of a Lipoprotein D fusion partner is included on the N-terminus to provide the polypeptide with additional exogenous T-cell epitopes and to increase the expression level in *E. coli* (thus functioning as an expression enhancer). The lipid tail ensures optimal presentation of the antigen to antigen presenting cells. Other fusion partners include the non-structural protein from influenzae virus, NS1 (hemaglutinin). Typically, the N-terminal 81 amino acids are used, although different fragments that include T-helper epitopes may be used.

In another embodiment, the immunological fusion partner is the protein known as LYTA, or a portion thereof (preferably a C-terminal portion). LYTA is derived from *Streptococcus pneumoniae*, which synthesizes an N-acetyl-L-alanine amidase known as amidase LYTA (encoded by the LytA gene; *Gene 43*:265-292, 1986). LYTA is an autolysin that specifically degrades certain bonds in the peptidoglycan backbone. The C-terminal domain of the LYTA protein is responsible for the affinity to the choline or to some choline analogues such as DEAE. This property has been exploited for the development of *E. coli* C-LYTA expressing plasmids useful for expression of fusion proteins. Purification of hybrid proteins containing the C-LYTA fragment at the amino terminus has been described (*see*

Biotechnology 10:795-798, 1992). Within a preferred embodiment, a repeat portion of LYTA may be incorporated into a fusion polypeptide. A repeat portion is found in the C-terminal region starting at residue 178. A particularly preferred repeat portion incorporates residues 188-305.

Yet another illustrative embodiment involves fusion polypeptides, and the polynucleotides encoding them, wherein the fusion partner comprises a targeting signal capable of directing a polypeptide to the endosomal/lysosomal compartment, as described in U.S. Patent No. 5,633,234. An immunogenic polypeptide of the invention, when fused with this targeting signal, will associate more efficiently with MHC class II molecules and thereby provide enhanced in vivo stimulation of $CD4^+$ T-cells specific for the polypeptide.

Polypeptides of the invention are prepared using any of a variety of well known synthetic and/or recombinant techniques, the latter of which are further described below. Polypeptides, portions and other variants generally less than about 150 amino acids can be generated by synthetic means, using techniques well known to those of ordinary skill in the art. In one illustrative example, such polypeptides are synthesized using any of the commercially available solid-phase techniques, such as the Merrifield solid-phase synthesis method, where amino acids are sequentially added to a growing amino acid chain. *See* Merrifield, *J. Am. Chem. Soc. 85*:2149-2146, 1963. Equipment for automated synthesis of polypeptides is commercially available from suppliers such as Perkin Elmer/Applied BioSystems Division (Foster City, CA), and may be operated according to the manufacturer's instructions.

In general, polypeptide compositions (including fusion polypeptides) of the invention are isolated. An "isolated" polypeptide is one that is removed from its original environment. For example, a naturally-occurring protein or polypeptide is isolated if it is separated from some or all of the coexisting materials in the natural system. Preferably, such polypeptides are also purified, *e.g.*, are at least about 90% pure, more preferably at least about 95% pure and most preferably at least about 99% pure.

POLYNUCLEOTIDE COMPOSITIONS

The present invention, in other aspects, provides polynucleotide compositions. The terms "DNA" and "polynucleotide" are used essentially interchangeably herein to refer to a DNA molecule that has been isolated free of total genomic DNA of a particular species. "Isolated," as used herein, means that a polynucleotide is substantially away from other coding sequences, and that the DNA molecule does not contain large portions of unrelated coding DNA, such as large chromosomal fragments or other functional genes or polypeptide coding regions. Of course, this refers to the DNA molecule as originally isolated, and does not exclude genes or coding regions later added to the segment by the hand of man.

As will be understood by those skilled in the art, the polynucleotide compositions of this invention can include genomic sequences, extra-genomic and plasmid-encoded sequences and smaller engineered gene segments that express, or may be adapted to express, proteins, polypeptides, peptides and the like. Such segments may be naturally isolated, or modified synthetically by the hand of man.

As will be also recognized by the skilled artisan, polynucleotides of the invention may be single-stranded (coding or antisense) or double-stranded, and may be DNA (genomic, cDNA or synthetic) or RNA molecules. RNA molecules may include HnRNA molecules, which contain introns and correspond to a DNA molecule in a one-to-one manner, and mRNA molecules, which do not contain introns. Additional coding or non-coding sequences may, but need not, be present within a polynucleotide of the present invention, and a polynucleotide may, but need not, be linked to other molecules and/or support materials.

Polynucleotides may comprise a native sequence (*i.e.*, an endogenous sequence that encodes a polypeptide/protein of the invention or a portion thereof) or may comprise a sequence that encodes a variant or derivative, preferably and immunogenic variant or derivative, of such a sequence.

Therefore, according to another aspect of the present invention, polynucleotide compositions are provided that comprise some or all of a polynucleotide sequence set forth in any one of SEQ ID NO:1-377, complements of a polynucleotide sequence set forth in any one of SEQ ID NO:1-377, and degenerate variants of a

polynucleotide sequence set forth in any one of SEQ ID NO:1-377. In certain preferred embodiments, the polynucleotide sequences set forth herein encode immunogenic polypeptides, as described above.

In other related embodiments, the present invention provides polynucleotide variants having substantial identity to the sequences disclosed herein in SEQ ID NO:1-377, for example those comprising at least 70% sequence identity, preferably at least 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% or higher, sequence identity compared to a polynucleotide sequence of this invention using the methods described herein, (*e.g.*, BLAST analysis using standard parameters, as described below). One skilled in this art will recognize that these values can be appropriately adjusted to determine corresponding identity of proteins encoded by two nucleotide sequences by taking into account codon degeneracy, amino acid similarity, reading frame positioning and the like.

Typically, polynucleotide variants will contain one or more substitutions, additions, deletions and/or insertions, preferably such that the immunogenicity of the polypeptide encoded by the variant polynucleotide is not substantially diminished relative to a polypeptide encoded by a polynucleotide sequence specifically set forth herein). The term "variants" should also be understood to encompasses homologous genes of xenogeneic origin.

In additional embodiments, the present invention provides polynucleotide fragments comprising or consisting of various lengths of contiguous stretches of sequence identical to or complementary to one or more of the sequences disclosed herein. For example, polynucleotides are provided by this invention that comprise or consist of at least about 10, 15, 20, 30, 40, 50, 75, 100, 150, 200, 300, 400, 500 or 1000 or more contiguous nucleotides of one or more of the sequences disclosed herein as well as all intermediate lengths there between. It will be readily understood that "intermediate lengths", in this context, means any length between the quoted values, such as 16, 17, 18, 19, *etc.*; 21, 22, 23, *etc.*; 30, 31, 32, *etc.*; 50, 51, 52, 53, *etc.*; 100, 101, 102, 103, *etc.*; 150, 151, 152, 153, *etc.*; including all integers through 200-500; 500-1,000, and the like. A polynucleotide sequence as described here may be extended at one or both ends by additional nucleotides not found in the native sequence.

This additional sequence may consist of 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20 nucleotides at either end of the disclosed sequence or at both ends of the disclosed sequence.

In another embodiment of the invention, polynucleotide compositions are provided that are capable of hybridizing under moderate to high stringency conditions to a polynucleotide sequence provided herein, or a fragment thereof, or a complementary sequence thereof. Hybridization techniques are well known in the art of molecular biology. For purposes of illustration, suitable moderately stringent conditions for testing the hybridization of a polynucleotide of this invention with other polynucleotides include prewashing in a solution of 5 X SSC, 0.5% SDS, 1.0 mM EDTA (pH 8.0); hybridizing at 50°C-60°C, 5 X SSC, overnight; followed by washing twice at 65°C for 20 minutes with each of 2X, 0.5X and 0.2X SSC containing 0.1% SDS. One skilled in the art will understand that the stringency of hybridization can be readily manipulated, such as by altering the salt content of the hybridization solution and/or the temperature at which the hybridization is performed. For example, in another embodiment, suitable highly stringent hybridization conditions include those described above, with the exception that the temperature of hybridization is increased, *e.g.*, to 60-65°C or 65-70°C.

In certain preferred embodiments, the polynucleotides described above, *e.g.*, polynucleotide variants, fragments and hybridizing sequences, encode polypeptides that are immunologically cross-reactive with a polypeptide sequence specifically set forth herein. In other preferred embodiments, such polynucleotides encode polypeptides that have a level of immunogenic activity of at least about 50%, preferably at least about 70%, and more preferably at least about 90% of that for a polypeptide sequence specifically set forth herein.

The polynucleotides of the present invention, or fragments thereof, regardless of the length of the coding sequence itself, may be combined with other DNA sequences, such as promoters, polyadenylation signals, additional restriction enzyme sites, multiple cloning sites, other coding segments, and the like, such that their overall length may vary considerably. It is therefore contemplated that a nucleic acid fragment of almost any length may be employed, with the total length preferably being

limited by the ease of preparation and use in the intended recombinant DNA protocol. For example, illustrative polynucleotide segments with total lengths of about 10,000, about 5000, about 3000, about 2,000, about 1,000, about 500, about 200, about 100, about 50 base pairs in length, and the like, (including all intermediate lengths) are contemplated to be useful in many implementations of this invention.

When comparing polynucleotide sequences, two sequences are said to be "identical" if the sequence of nucleotides in the two sequences is the same when aligned for maximum correspondence, as described below. Comparisons between two sequences are typically performed by comparing the sequences over a comparison window to identify and compare local regions of sequence similarity. A "comparison window" as used herein, refers to a segment of at least about 20 contiguous positions, usually 30 to about 75, 40 to about 50, in which a sequence may be compared to a reference sequence of the same number of contiguous positions after the two sequences are optimally aligned.

Optimal alignment of sequences for comparison may be conducted using the Megalign program in the Lasergene suite of bioinformatics software (DNASTAR, Inc., Madison, WI), using default parameters. This program embodies several alignment schemes described in the following references: Dayhoff, M.O. (1978) A model of evolutionary change in proteins – Matrices for detecting distant relationships. In Dayhoff, M.O. (ed.) Atlas of Protein Sequence and Structure, National Biomedical Research Foundation, Washington DC Vol. 5, Suppl. 3, pp. 345-358; Hein J. (1990) Unified Approach to Alignment and Phylogenes pp. 626-645 *Methods in Enzymology* vol. 183, Academic Press, Inc., San Diego, CA; Higgins, D.G. and Sharp, P.M. (1989) *CABIOS 5*:151-153; Myers, E.W. and Muller W. (1988) *CABIOS 4*:11-17; Robinson, E.D. (1971) *Comb. Theor 11*:105; Santou, N. Nes, M. (1987) *Mol. Biol. Evol. 4*:406-425; Sneath, P.H.A. and Sokal, R.R. (1973) *Numerical Taxonomy – the Principles and Practice of Numerical Taxonomy*, Freeman Press, San Francisco, CA; Wilbur, W.J. and Lipman, D.J. (1983) *Proc. Natl. Acad., Sci. USA 80*:726-730.

Alternatively, optimal alignment of sequences for comparison may be conducted by the local identity algorithm of Smith and Waterman (1981) *Add. APL. Math* 2:482, by the identity alignment algorithm of Needleman and Wunsch (1970) *J.*

Mol. Biol. 48:443, by the search for similarity methods of Pearson and Lipman (1988) *Proc. Natl. Acad. Sci. USA* 85: 2444, by computerized implementations of these algorithms (GAP, BESTFIT, BLAST, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group (GCG), 575 Science Dr., Madison, WI), or by inspection.

One preferred example of algorithms that are suitable for determining percent sequence identity and sequence similarity are the BLAST and BLAST 2.0 algorithms, which are described in Altschul et al. (1977) *Nucl. Acids Res.* 25:3389-3402 and Altschul et al. (1990) *J. Mol. Biol.* 215:403-410, respectively. BLAST and BLAST 2.0 can be used, for example with the parameters described herein, to determine percent sequence identity for the polynucleotides of the invention. Software for performing BLAST analyses is publicly available through the National Center for Biotechnology Information. In one illustrative example, cumulative scores can be calculated using, for nucleotide sequences, the parameters M (reward score for a pair of matching residues; always >0) and N (penalty score for mismatching residues; always <0). Extension of the word hits in each direction are halted when: the cumulative alignment score falls off by the quantity X from its maximum achieved value; the cumulative score goes to zero or below, due to the accumulation of one or more negative-scoring residue alignments; or the end of either sequence is reached. The BLAST algorithm parameters W, T and X determine the sensitivity and speed of the alignment. The BLASTN program (for nucleotide sequences) uses as defaults a wordlength (W) of 11, and expectation (E) of 10, and the BLOSUM62 scoring matrix (see Henikoff and Henikoff (1989) *Proc. Natl. Acad. Sci. USA* 89:10915) alignments, (B) of 50, expectation (E) of 10, M=5, N=-4 and a comparison of both strands.

Preferably, the "percentage of sequence identity" is determined by comparing two optimally aligned sequences over a window of comparison of at least 20 positions, wherein the portion of the polynucleotide sequence in the comparison window may comprise additions or deletions (*i.e.*, gaps) of 20 percent or less, usually 5 to 15 percent, or 10 to 12 percent, as compared to the reference sequences (which does not comprise additions or deletions) for optimal alignment of the two sequences. The percentage is calculated by determining the number of positions at which the identical

nucleic acid bases occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the reference sequence (*i.e.,* the window size) and multiplying the results by 100 to yield the percentage of sequence identity.

It will be appreciated by those of ordinary skill in the art that, as a result of the degeneracy of the genetic code, there are many nucleotide sequences that encode a polypeptide as described herein. Some of these polynucleotides bear minimal homology to the nucleotide sequence of any native gene. Nonetheless, polynucleotides that vary due to differences in codon usage are specifically contemplated by the present invention. Further, alleles of the genes comprising the polynucleotide sequences provided herein are within the scope of the present invention. Alleles are endogenous genes that are altered as a result of one or more mutations, such as deletions, additions and/or substitutions of nucleotides. The resulting mRNA and protein may, but need not, have an altered structure or function. Alleles may be identified using standard techniques (such as hybridization, amplification and/or database sequence comparison).

Therefore, in another embodiment of the invention, a mutagenesis approach, such as site-specific mutagenesis, is employed for the preparation of immunogenic variants and/or derivatives of the polypeptides described herein. By this approach, specific modifications in a polypeptide sequence can be made through mutagenesis of the underlying polynucleotides that encode them. These techniques provides a straightforward approach to prepare and test sequence variants, for example, incorporating one or more of the foregoing considerations, by introducing one or more nucleotide sequence changes into the polynucleotide.

Site-specific mutagenesis allows the production of mutants through the use of specific oligonucleotide sequences which encode the DNA sequence of the desired mutation, as well as a sufficient number of adjacent nucleotides, to provide a primer sequence of sufficient size and sequence complexity to form a stable duplex on both sides of the deletion junction being traversed. Mutations may be employed in a selected polynucleotide sequence to improve, alter, decrease, modify, or otherwise change the properties of the polynucleotide itself, and/or alter the properties, activity, composition, stability, or primary sequence of the encoded polypeptide.

In certain embodiments of the present invention, the inventors contemplate the mutagenesis of the disclosed polynucleotide sequences to alter one or more properties of the encoded polypeptide, such as the immunogenicity of a polypeptide vaccine. The techniques of site-specific mutagenesis are well known in the art, and are widely used to create variants of both polypeptides and polynucleotides. For example, site-specific mutagenesis is often used to alter a specific portion of a DNA molecule. In such embodiments, a primer comprising typically about 14 to about 25 nucleotides or so in length is employed, with about 5 to about 10 residues on both sides of the junction of the sequence being altered.

As will be appreciated by those of skill in the art, site-specific mutagenesis techniques have often employed a phage vector that exists in both a single stranded and double stranded form. Typical vectors useful in site-directed mutagenesis include vectors such as the M13 phage. These phage are readily commercially-available and their use is generally well-known to those skilled in the art. Double-stranded plasmids are also routinely employed in site directed mutagenesis that eliminates the step of transferring the gene of interest from a plasmid to a phage.

In general, site-directed mutagenesis in accordance herewith is performed by first obtaining a single-stranded vector or melting apart of two strands of a double-stranded vector that includes within its sequence a DNA sequence that encodes the desired peptide. An oligonucleotide primer bearing the desired mutated sequence is prepared, generally synthetically. This primer is then annealed with the single-stranded vector, and subjected to DNA polymerizing enzymes such as *E. coli* polymerase I Klenow fragment, in order to complete the synthesis of the mutation-bearing strand. Thus, a heteroduplex is formed wherein one strand encodes the original non-mutated sequence and the second strand bears the desired mutation. This heteroduplex vector is then used to transform appropriate cells, such as *E. coli* cells, and clones are selected which include recombinant vectors bearing the mutated sequence arrangement.

The preparation of sequence variants of the selected peptide-encoding DNA segments using site-directed mutagenesis provides a means of producing potentially useful species and is not meant to be limiting as there are other ways in

which sequence variants of peptides and the DNA sequences encoding them may be obtained. For example, recombinant vectors encoding the desired peptide sequence may be treated with mutagenic agents, such as hydroxylamine, to obtain sequence variants. Specific details regarding these methods and protocols are found in the teachings of Maloy *et al.*, 1994; Segal, 1976; Prokop and Bajpai, 1991; Kuby, 1994; and Maniatis *et al.*, 1982, each incorporated herein by reference, for that purpose.

As used herein, the term "oligonucleotide directed mutagenesis procedure" refers to template-dependent processes and vector-mediated propagation which result in an increase in the concentration of a specific nucleic acid molecule relative to its initial concentration, or in an increase in the concentration of a detectable signal, such as amplification. As used herein, the term "oligonucleotide directed mutagenesis procedure" is intended to refer to a process that involves the template-dependent extension of a primer molecule. The term template dependent process refers to nucleic acid synthesis of an RNA or a DNA molecule wherein the sequence of the newly synthesized strand of nucleic acid is dictated by the well-known rules of complementary base pairing (see, for example, Watson, 1987). Typically, vector mediated methodologies involve the introduction of the nucleic acid fragment into a DNA or RNA vector, the clonal amplification of the vector, and the recovery of the amplified nucleic acid fragment. Examples of such methodologies are provided by U. S. Patent No. 4,237,224, specifically incorporated herein by reference in its entirety.

In another approach for the production of polypeptide variants of the present invention, recursive sequence recombination, as described in U.S. Patent No. 5,837,458, may be employed. In this approach, iterative cycles of recombination and screening or selection are performed to "evolve" individual polynucleotide variants of the invention having, for example, enhanced immunogenic activity.

In other embodiments of the present invention, the polynucleotide sequences provided herein can be advantageously used as probes or primers for nucleic acid hybridization. As such, it is contemplated that nucleic acid segments that comprise or consist of a sequence region of at least about a 15 nucleotide long contiguous sequence that has the same sequence as, or is complementary to, a 15 nucleotide long contiguous sequence disclosed herein will find particular utility. Longer contiguous

identical or complementary sequences, *e.g.*, those of about 20, 30, 40, 50, 100, 200, 500, 1000 (including all intermediate lengths) and even up to full length sequences will also be of use in certain embodiments.

The ability of such nucleic acid probes to specifically hybridize to a sequence of interest will enable them to be of use in detecting the presence of complementary sequences in a given sample. However, other uses are also envisioned, such as the use of the sequence information for the preparation of mutant species primers, or primers for use in preparing other genetic constructions.

Polynucleotide molecules having sequence regions consisting of contiguous nucleotide stretches of 10-14, 15-20, 30, 50, or even of 100-200 nucleotides or so (including intermediate lengths as well), identical or complementary to a polynucleotide sequence disclosed herein, are particularly contemplated as hybridization probes for use in, *e.g.*, Southern and Northern blotting. This would allow a gene product, or fragment thereof, to be analyzed, both in diverse cell types and also in various bacterial cells. The total size of fragment, as well as the size of the complementary stretch(es), will ultimately depend on the intended use or application of the particular nucleic acid segment. Smaller fragments will generally find use in hybridization embodiments, wherein the length of the contiguous complementary region may be varied, such as between about 15 and about 100 nucleotides, but larger contiguous complementarity stretches may be used, according to the length complementary sequences one wishes to detect.

The use of a hybridization probe of about 15-25 nucleotides in length allows the formation of a duplex molecule that is both stable and selective. Molecules having contiguous complementary sequences over stretches greater than 15 bases in length are generally preferred, though, in order to increase stability and selectivity of the hybrid, and thereby improve the quality and degree of specific hybrid molecules obtained. One will generally prefer to design nucleic acid molecules having gene-complementary stretches of 15 to 25 contiguous nucleotides, or even longer where desired.

Hybridization probes may be selected from any portion of any of the sequences disclosed herein. All that is required is to review the sequences set forth

herein, or to any continuous portion of the sequences, from about 15-25 nucleotides in length up to and including the full length sequence, that one wishes to utilize as a probe or primer. The choice of probe and primer sequences may be governed by various factors. For example, one may wish to employ primers from towards the termini of the total sequence.

Small polynucleotide segments or fragments may be readily prepared by, for example, directly synthesizing the fragment by chemical means, as is commonly practiced using an automated oligonucleotide synthesizer. Also, fragments may be obtained by application of nucleic acid reproduction technology, such as the PCR™ technology of U. S. Patent 4,683,202 (incorporated herein by reference), by introducing selected sequences into recombinant vectors for recombinant production, and by other recombinant DNA techniques generally known to those of skill in the art of molecular biology.

The nucleotide sequences of the invention may be used for their ability to selectively form duplex molecules with complementary stretches of the entire gene or gene fragments of interest. Depending on the application envisioned, one will typically desire to employ varying conditions of hybridization to achieve varying degrees of selectivity of probe towards target sequence. For applications requiring high selectivity, one will typically desire to employ relatively stringent conditions to form the hybrids, *e.g.*, one will select relatively low salt and/or high temperature conditions, such as provided by a salt concentration of from about 0.02 M to about 0.15 M salt at temperatures of from about 50°C to about 70°C. Such selective conditions tolerate little, if any, mismatch between the probe and the template or target strand, and would be particularly suitable for isolating related sequences.

Of course, for some applications, for example, where one desires to prepare mutants employing a mutant primer strand hybridized to an underlying template, less stringent (reduced stringency) hybridization conditions will typically be needed in order to allow formation of the heteroduplex. In these circumstances, one may desire to employ salt conditions such as those of from about 0.15 M to about 0.9 M salt, at temperatures ranging from about 20°C to about 55°C. Cross-hybridizing species can thereby be readily identified as positively hybridizing signals with respect to control

hybridizations. In any case, it is generally appreciated that conditions can be rendered more stringent by the addition of increasing amounts of formamide, which serves to destabilize the hybrid duplex in the same manner as increased temperature. Thus, hybridization conditions can be readily manipulated, and thus will generally be a method of choice depending on the desired results.

According to another embodiment of the present invention, polynucleotide compositions comprising antisense oligonucleotides are provided. Antisense oligonucleotides have been demonstrated to be effective and targeted inhibitors of protein synthesis, and, consequently, provide a therapeutic approach by which a disease can be treated by inhibiting the synthesis of proteins that contribute to the disease. The efficacy of antisense oligonucleotides for inhibiting protein synthesis is well established. For example, the synthesis of polygalactauronase and the muscarine type 2 acetylcholine receptor are inhibited by antisense oligonucleotides directed to their respective mRNA sequences (U. S. Patent 5,739,119 and U. S. Patent 5,759,829). Further, examples of antisense inhibition have been demonstrated with the nuclear protein cyclin, the multiple drug resistance gene (MDG1), ICAM-1, E-selectin, STK-1, striatal $GABA_A$ receptor and human EGF (Jaskulski *et al.*, Science. 1988 Jun 10;240(4858):1544-6; Vasanthakumar and Ahmed, Cancer Commun. 1989;1(4):225-32; Peris *et al.*, Brain Res Mol Brain Res. 1998 Jun 15;57(2):310-20; U. S. Patent 5,801,154; U.S. Patent 5,789,573; U. S. Patent 5,718,709 and U.S. Patent 5,610,288). Antisense constructs have also been described that inhibit and can be used to treat a variety of abnormal cellular proliferations, *e.g.* cancer (U. S. Patent 5,747,470; U. S. Patent 5,591,317 and U. S. Patent 5,783,683).

Therefore, in certain embodiments, the present invention provides oligonucleotide sequences that comprise all, or a portion of, any sequence that is capable of specifically binding to polynucleotide sequence described herein, or a complement thereof. In one embodiment, the antisense oligonucleotides comprise DNA or derivatives thereof. In another embodiment, the oligonucleotides comprise RNA or derivatives thereof. In a third embodiment, the oligonucleotides are modified DNAs comprising a phosphorothioated modified backbone. In a fourth embodiment, the oligonucleotide sequences comprise peptide nucleic acids or derivatives thereof. In

each case, preferred compositions comprise a sequence region that is complementary, and more preferably substantially-complementary, and even more preferably, completely complementary to one or more portions of polynucleotides disclosed herein. Selection of antisense compositions specific for a given gene sequence is based upon analysis of the chosen target sequence and determination of secondary structure, T_m, binding energy, and relative stability. Antisense compositions may be selected based upon their relative inability to form dimers, hairpins, or other secondary structures that would reduce or prohibit specific binding to the target mRNA in a host cell. Highly preferred target regions of the mRNA, are those which are at or near the AUG translation initiation codon, and those sequences which are substantially complementary to 5' regions of the mRNA. These secondary structure analyses and target site selection considerations can be performed, for example, using v.4 of the OLIGO primer analysis software and/or the BLASTN 2.0.5 algorithm software (Altschul *et al.*, Nucleic Acids Res. 1997, 25(17):3389-402).

The use of an antisense delivery method employing a short peptide vector, termed MPG (27 residues), is also contemplated. The MPG peptide contains a hydrophobic domain derived from the fusion sequence of HIV gp41 and a hydrophilic domain from the nuclear localization sequence of SV40 T-antigen (Morris *et al.*, Nucleic Acids Res. 1997 Jul 15;25(14):2730-6). It has been demonstrated that several molecules of the MPG peptide coat the antisense oligonucleotides and can be delivered into cultured mammalian cells in less than 1 hour with relatively high efficiency (90%). Further, the interaction with MPG strongly increases both the stability of the oligonucleotide to nuclease and the ability to cross the plasma membrane.

According to another embodiment of the invention, the polynucleotide compositions described herein are used in the design and preparation of ribozyme molecules for inhibiting expression of the tumor polypeptides and proteins of the present invention in tumor cells. Ribozymes are RNA-protein complexes that cleave nucleic acids in a site-specific fashion. Ribozymes have specific catalytic domains that possess endonuclease activity (Kim and Cech, Proc Natl Acad Sci U S A. 1987 Dec;84(24):8788-92; Forster and Symons, Cell. 1987 Apr 24;49(2):211-20). For example, a large number of ribozymes accelerate phosphoester transfer reactions with a

high degree of specificity, often cleaving only one of several phosphoesters in an oligonucleotide substrate (Cech *et al.*, Cell. 1981 Dec;27(3 Pt 2):487-96; Michel and Westhof, J Mol Biol. 1990 Dec 5;216(3):585-610; Reinhold-Hurek and Shub, Nature. 1992 May 14;357(6374):173-6). This specificity has been attributed to the requirement that the substrate bind via specific base-pairing interactions to the internal guide sequence ("IGS") of the ribozyme prior to chemical reaction.

Six basic varieties of naturally occurring enzymatic RNAs are known presently. Each can catalyze the hydrolysis of RNA phosphodiester bonds *in trans* (and thus can cleave other RNA molecules) under physiological conditions. In general, enzymatic nucleic acids act by first binding to a target RNA. Such binding occurs through the target binding portion of a enzymatic nucleic acid which is held in close proximity to an enzymatic portion of the molecule that acts to cleave the target RNA. Thus, the enzymatic nucleic acid first recognizes and then binds a target RNA through complementary base-pairing, and once bound to the correct site, acts enzymatically to cut the target RNA. Strategic cleavage of such a target RNA will destroy its ability to direct synthesis of an encoded protein. After an enzymatic nucleic acid has bound and cleaved its RNA target, it is released from that RNA to search for another target and can repeatedly bind and cleave new targets.

The enzymatic nature of a ribozyme is advantageous over many technologies, such as antisense technology (where a nucleic acid molecule simply binds to a nucleic acid target to block its translation) since the concentration of ribozyme necessary to affect a therapeutic treatment is lower than that of an antisense oligonucleotide. This advantage reflects the ability of the ribozyme to act enzymatically. Thus, a single ribozyme molecule is able to cleave many molecules of target RNA. In addition, the ribozyme is a highly specific inhibitor, with the specificity of inhibition depending not only on the base pairing mechanism of binding to the target RNA, but also on the mechanism of target RNA cleavage. Single mismatches, or base-substitutions, near the site of cleavage can completely eliminate catalytic activity of a ribozyme. Similar mismatches in antisense molecules do not prevent their action (Woolf *et al.*, Proc Natl Acad Sci U S A. 1992 Aug 15;89(16):7305-9). Thus, the

specificity of action of a ribozyme is greater than that of an antisense oligonucleotide binding the same RNA site.

The enzymatic nucleic acid molecule may be formed in a hammerhead, hairpin, a hepatitis δ virus, group I intron or RNaseP RNA (in association with an RNA guide sequence) or Neurospora VS RNA motif. Examples of hammerhead motifs are described by Rossi *et al.* Nucleic Acids Res. 1992 Sep 11;20(17):4559-65. Examples of hairpin motifs are described by Hampel *et al.* (Eur. Pat. Appl. Publ. No. EP 0360257), Hampel and Tritz, Biochemistry 1989 Jun 13;28(12):4929-33; Hampel *et al.,* Nucleic Acids Res. 1990 Jan 25;18(2):299-304 and U. S. Patent 5,631,359. An example of the hepatitis δ virus motif is described by Perrotta and Been, Biochemistry. 1992 Dec 1;31(47):11843-52; an example of the RNaseP motif is described by Guerrier-Takada *et al.*, Cell. 1983 Dec;35(3 Pt 2):849-57; Neurospora VS RNA ribozyme motif is described by Collins (Saville and Collins, Cell. 1990 May 18;61(4):685-96; Saville and Collins, Proc Natl Acad Sci U S A. 1991 Oct 1;88(19):8826-30; Collins and Olive, Biochemistry. 1993 Mar 23;32(11):2795-9); and an example of the Group I intron is described in (U. S. Patent 4,987,071). All that is important in an enzymatic nucleic acid molecule of this invention is that it has a specific substrate binding site which is complementary to one or more of the target gene RNA regions, and that it have nucleotide sequences within or surrounding that substrate binding site which impart an RNA cleaving activity to the molecule. Thus the ribozyme constructs need not be limited to specific motifs mentioned herein.

Ribozymes may be designed as described in Int. Pat. Appl. Publ. No. WO 93/23569 and Int. Pat. Appl. Publ. No. WO 94/02595, each specifically incorporated herein by reference) and synthesized to be tested *in vitro* and *in vivo*, as described. Such ribozymes can also be optimized for delivery. While specific examples are provided, those in the art will recognize that equivalent RNA targets in other species can be utilized when necessary.

Ribozyme activity can be optimized by altering the length of the ribozyme binding arms, or chemically synthesizing ribozymes with modifications that prevent their degradation by serum ribonucleases (see *e.g.*, Int. Pat. Appl. Publ. No. WO 92/07065; Int. Pat. Appl. Publ. No. WO 93/15187; Int. Pat. Appl. Publ. No. WO

91/03162; Eur. Pat. Appl. Publ. No. 92110298.4; U. S. Patent 5,334,711; and Int. Pat. Appl. Publ. No. WO 94/13688, which describe various chemical modifications that can be made to the sugar moieties of enzymatic RNA molecules), modifications which enhance their efficacy in cells, and removal of stem II bases to shorten RNA synthesis times and reduce chemical requirements.

Sullivan *et al.* (Int. Pat. Appl. Publ. No. WO 94/02595) describes the general methods for delivery of enzymatic RNA molecules. Ribozymes may be administered to cells by a variety of methods known to those familiar to the art, including, but not restricted to, encapsulation in liposomes, by iontophoresis, or by incorporation into other vehicles, such as hydrogels, cyclodextrins, biodegradable nanocapsules, and bioadhesive microspheres. For some indications, ribozymes may be directly delivered *ex vivo* to cells or tissues with or without the aforementioned vehicles. Alternatively, the RNA/vehicle combination may be locally delivered by direct inhalation, by direct injection or by use of a catheter, infusion pump or stent. Other routes of delivery include, but are not limited to, intravascular, intramuscular, subcutaneous or joint injection, aerosol inhalation, oral (tablet or pill form), topical, systemic, ocular, intraperitoneal and/or intrathecal delivery. More detailed descriptions of ribozyme delivery and administration are provided in Int. Pat. Appl. Publ. No. WO 94/02595 and Int. Pat. Appl. Publ. No. WO 93/23569, each specifically incorporated herein by reference.

Another means of accumulating high concentrations of a ribozyme(s) within cells is to incorporate the ribozyme-encoding sequences into a DNA expression vector. Transcription of the ribozyme sequences are driven from a promoter for eukaryotic RNA polymerase I (pol I), RNA polymerase II (pol II), or RNA polymerase III (pol III). Transcripts from pol II or pol III promoters will be expressed at high levels in all cells; the levels of a given pol II promoter in a given cell type will depend on the nature of the gene regulatory sequences (enhancers, silencers, *etc.*) present nearby. Prokaryotic RNA polymerase promoters may also be used, providing that the prokaryotic RNA polymerase enzyme is expressed in the appropriate cells Ribozymes expressed from such promoters have been shown to function in mammalian cells. Such transcription units can be incorporated into a variety of vectors for introduction into

mammalian cells, including but not restricted to, plasmid DNA vectors, viral DNA vectors (such as adenovirus or adeno-associated vectors), or viral RNA vectors (such as retroviral, semliki forest virus, sindbis virus vectors).

In another embodiment of the invention, peptide nucleic acids (PNAs) compositions are provided. PNA is a DNA mimic in which the nucleobases are attached to a pseudopeptide backbone (Good and Nielsen, Antisense Nucleic Acid Drug Dev. 1997 7(4) 431-37). PNA is able to be utilized in a number methods that traditionally have used RNA or DNA. Often PNA sequences perform better in techniques than the corresponding RNA or DNA sequences and have utilities that are not inherent to RNA or DNA. A review of PNA including methods of making, characteristics of, and methods of using, is provided by Corey (*Trends Biotechnol* 1997 Jun;15(6):224-9). As such, in certain embodiments, one may prepare PNA sequences that are complementary to one or more portions of the ACE mRNA sequence, and such PNA compositions may be used to regulate, alter, decrease, or reduce the translation of ACE-specific mRNA, and thereby alter the level of ACE activity in a host cell to which such PNA compositions have been administered.

PNAs have 2-aminoethyl-glycine linkages replacing the normal phosphodiester backbone of DNA (Nielsen *et al.*, *Science* 1991 Dec 6;254(5037):1497-500; Hanvey *et al.*, Science. 1992 Nov 27;258(5087):1481-5; Hyrup and Nielsen, Bioorg Med Chem. 1996 Jan;4(1):5-23). This chemistry has three important consequences: firstly, in contrast to DNA or phosphorothioate oligonucleotides, PNAs are neutral molecules; secondly, PNAs are achiral, which avoids the need to develop a stereoselective synthesis; and thirdly, PNA synthesis uses standard Boc or Fmoc protocols for solid-phase peptide synthesis, although other methods, including a modified Merrifield method, have been used.

PNA monomers or ready-made oligomers are commercially available from PerSeptive Biosystems (Framingham, MA). PNA syntheses by either Boc or Fmoc protocols are straightforward using manual or automated protocols (Norton *et al.*, Bioorg Med Chem. 1995 Apr;3(4):437-45). The manual protocol lends itself to the production of chemically modified PNAs or the simultaneous synthesis of families of closely related PNAs.

As with peptide synthesis, the success of a particular PNA synthesis will depend on the properties of the chosen sequence. For example, while in theory PNAs can incorporate any combination of nucleotide bases, the presence of adjacent purines can lead to deletions of one or more residues in the product. In expectation of this difficulty, it is suggested that, in producing PNAs with adjacent purines, one should repeat the coupling of residues likely to be added inefficiently. This should be followed by the purification of PNAs by reverse-phase high-pressure liquid chromatography, providing yields and purity of product similar to those observed during the synthesis of peptides.

Modifications of PNAs for a given application may be accomplished by coupling amino acids during solid-phase synthesis or by attaching compounds that contain a carboxylic acid group to the exposed N-terminal amine. Alternatively, PNAs can be modified after synthesis by coupling to an introduced lysine or cysteine. The ease with which PNAs can be modified facilitates optimization for better solubility or for specific functional requirements. Once synthesized, the identity of PNAs and their derivatives can be confirmed by mass spectrometry. Several studies have made and utilized modifications of PNAs (for example, Norton *et al.*, Bioorg Med Chem. 1995 Apr;3(4):437-45; Petersen *et al.*, J Pept Sci. 1995 May-Jun;1(3):175-83; Orum *et al.*, Biotechniques. 1995 Sep;19(3):472-80; Footer *et al.*, Biochemistry. 1996 Aug 20;35(33):10673-9; Griffith *et al.*, Nucleic Acids Res. 1995 Aug 11;23(15):3003-8; Pardridge *et al.*, Proc Natl Acad Sci U S A. 1995 Jun 6;92(12):5592-6; Boffa *et al.*, Proc Natl Acad Sci U S A. 1995 Mar 14;92(6):1901-5; Gambacorti-Passerini *et al.*, Blood. 1996 Aug 15;88(4):1411-7; Armitage *et al.*, Proc Natl Acad Sci U S A. 1997 Nov 11;94(23):12320-5; Seeger *et al.*, Biotechniques. 1997 Sep;23(3):512-7). U.S. Patent No. 5,700,922 discusses PNA-DNA-PNA chimeric molecules and their uses in diagnostics, modulating protein in organisms, and treatment of conditions susceptible to therapeutics.

Methods of characterizing the antisense binding properties of PNAs are discussed in Rose (Anal Chem. 1993 Dec 15;65(24):3545-9) and Jensen *et al.* (Biochemistry. 1997 Apr 22;36(16):5072-7). Rose uses capillary gel electrophoresis to determine binding of PNAs to their complementary oligonucleotide, measuring the

relative binding kinetics and stoichiometry. Similar types of measurements were made by Jensen *et al.* using BIAcore™ technology.

Other applications of PNAs that have been described and will be apparent to the skilled artisan include use in DNA strand invasion, antisense inhibition, mutational analysis, enhancers of transcription, nucleic acid purification, isolation of transcriptionally active genes, blocking of transcription factor binding, genome cleavage, biosensors, *in situ* hybridization, and the like.

POLYNUCLEOTIDE IDENTIFICATION, CHARACTERIZATION AND EXPRESSION

Polynucleotides compositions of the present invention may be identified, prepared and/or manipulated using any of a variety of well established techniques (see generally, Sambrook et al., *Molecular Cloning: A Laboratory Manual,* Cold Spring Harbor Laboratories, Cold Spring Harbor, NY, 1989, and other like references). For example, a polynucleotide may be identified, as described in more detail below, by screening a microarray of cDNAs for tumor-associated expression (*i.e.,* expression that is at least two fold greater in a tumor than in normal tissue, as determined using a representative assay provided herein). Such screens may be performed, for example, using the microarray technology of Affymetrix, Inc. (Santa Clara, CA) according to the manufacturer's instructions (and essentially as described by Schena et al., *Proc. Natl. Acad. Sci. USA 93*:10614-10619, 1996 and Heller et al., *Proc. Natl. Acad. Sci. USA 94*:2150-2155, 1997). Alternatively, polynucleotides may be amplified from cDNA prepared from cells expressing the proteins described herein, such as tumor cells.

Many template dependent processes are available to amplify a target sequences of interest present in a sample. One of the best known amplification methods is the polymerase chain reaction (PCR™) which is described in detail in U.S. Patent Nos. 4,683,195, 4,683,202 and 4,800,159, each of which is incorporated herein by reference in its entirety. Briefly, in PCR™, two primer sequences are prepared which are complementary to regions on opposite complementary strands of the target sequence. An excess of deoxynucleoside triphosphates is added to a reaction mixture along with a DNA polymerase (*e.g.*, *Taq* polymerase). If the target sequence is present in a sample, the primers will bind to the target and the polymerase will cause the

primers to be extended along the target sequence by adding on nucleotides. By raising and lowering the temperature of the reaction mixture, the extended primers will dissociate from the target to form reaction products, excess primers will bind to the target and to the reaction product and the process is repeated. Preferably reverse transcription and PCR™ amplification procedure may be performed in order to quantify the amount of mRNA amplified. Polymerase chain reaction methodologies are well known in the art.

Any of a number of other template dependent processes, many of which are variations of the PCR ™ amplification technique, are readily known and available in the art. Illustratively, some such methods include the ligase chain reaction (referred to as LCR), described, for example, in Eur. Pat. Appl. Publ. No. 320,308 and U.S. Patent No. 4,883,750; Qbeta Replicase, described in PCT Intl. Pat. Appl. Publ. No. PCT/US87/00880; Strand Displacement Amplification (SDA) and Repair Chain Reaction (RCR). Still other amplification methods are described in Great Britain Pat. Appl. No. 2 202 328, and in PCT Intl. Pat. Appl. Publ. No. PCT/US89/01025. Other nucleic acid amplification procedures include transcription-based amplification systems (TAS) (PCT Intl. Pat. Appl. Publ. No. WO 88/10315), including nucleic acid sequence based amplification (NASBA) and 3SR. Eur. Pat. Appl. Publ. No. 329,822 describes a nucleic acid amplification process involving cyclically synthesizing single-stranded RNA ("ssRNA"), ssDNA, and double-stranded DNA (dsDNA). PCT Intl. Pat. Appl. Publ. No. WO 89/06700 describes a nucleic acid sequence amplification scheme based on the hybridization of a promoter/primer sequence to a target single-stranded DNA ("ssDNA") followed by transcription of many RNA copies of the sequence. Other amplification methods such as "RACE" (Frohman, 1990), and "one-sided PCR" (Ohara, 1989) are also well-known to those of skill in the art.

An amplified portion of a polynucleotide of the present invention may be used to isolate a full length gene from a suitable library (*e.g.,* a tumor cDNA library) using well known techniques. Within such techniques, a library (cDNA or genomic) is screened using one or more polynucleotide probes or primers suitable for amplification. Preferably, a library is size-selected to include larger molecules. Random primed

libraries may also be preferred for identifying 5' and upstream regions of genes. Genomic libraries are preferred for obtaining introns and extending 5' sequences.

For hybridization techniques, a partial sequence may be labeled (*e.g.*, by nick-translation or end-labeling with ^{32}P) using well known techniques. A bacterial or bacteriophage library is then generally screened by hybridizing filters containing denatured bacterial colonies (or lawns containing phage plaques) with the labeled probe (*see* Sambrook et al., *Molecular Cloning: A Laboratory Manual,* Cold Spring Harbor Laboratories, Cold Spring Harbor, NY, 1989). Hybridizing colonies or plaques are selected and expanded, and the DNA is isolated for further analysis. cDNA clones may be analyzed to determine the amount of additional sequence by, for example, PCR using a primer from the partial sequence and a primer from the vector. Restriction maps and partial sequences may be generated to identify one or more overlapping clones. The complete sequence may then be determined using standard techniques, which may involve generating a series of deletion clones. The resulting overlapping sequences can then assembled into a single contiguous sequence. A full length cDNA molecule can be generated by ligating suitable fragments, using well known techniques.

Alternatively, amplification techniques, such as those described above, can be useful for obtaining a full length coding sequence from a partial cDNA sequence. One such amplification technique is inverse PCR (*see* Triglia et al., *Nucl. Acids Res. 16*:8186, 1988), which uses restriction enzymes to generate a fragment in the known region of the gene. The fragment is then circularized by intramolecular ligation and used as a template for PCR with divergent primers derived from the known region. Within an alternative approach, sequences adjacent to a partial sequence may be retrieved by amplification with a primer to a linker sequence and a primer specific to a known region. The amplified sequences are typically subjected to a second round of amplification with the same linker primer and a second primer specific to the known region. A variation on this procedure, which employs two primers that initiate extension in opposite directions from the known sequence, is described in WO 96/38591. Another such technique is known as "rapid amplification of cDNA ends" or RACE. This technique involves the use of an internal primer and an external primer, which hybridizes to a polyA region or vector sequence, to identify sequences that are 5'

and 3' of a known sequence. Additional techniques include capture PCR (Lagerstrom et al., *PCR Methods Applic. 1*:111-19, 1991) and walking PCR (Parker et al., *Nucl. Acids. Res. 19*:3055-60, 1991). Other methods employing amplification may also be employed to obtain a full length cDNA sequence.

In certain instances, it is possible to obtain a full length cDNA sequence by analysis of sequences provided in an expressed sequence tag (EST) database, such as that available from GenBank. Searches for overlapping ESTs may generally be performed using well known programs (*e.g.,* NCBI BLAST searches), and such ESTs may be used to generate a contiguous full length sequence. Full length DNA sequences may also be obtained by analysis of genomic fragments.

In other embodiments of the invention, polynucleotide sequences or fragments thereof which encode polypeptides of the invention, or fusion proteins or functional equivalents thereof, may be used in recombinant DNA molecules to direct expression of a polypeptide in appropriate host cells. Due to the inherent degeneracy of the genetic code, other DNA sequences that encode substantially the same or a functionally equivalent amino acid sequence may be produced and these sequences may be used to clone and express a given polypeptide.

As will be understood by those of skill in the art, it may be advantageous in some instances to produce polypeptide-encoding nucleotide sequences possessing non-naturally occurring codons. For example, codons preferred by a particular prokaryotic or eukaryotic host can be selected to increase the rate of protein expression or to produce a recombinant RNA transcript having desirable properties, such as a half-life which is longer than that of a transcript generated from the naturally occurring sequence.

Moreover, the polynucleotide sequences of the present invention can be engineered using methods generally known in the art in order to alter polypeptide encoding sequences for a variety of reasons, including but not limited to, alterations which modify the cloning, processing, and/or expression of the gene product. For example, DNA shuffling by random fragmentation and PCR reassembly of gene fragments and synthetic oligonucleotides may be used to engineer the nucleotide sequences. In addition, site-directed mutagenesis may be used to insert new restriction

sites, alter glycosylation patterns, change codon preference, produce splice variants, or introduce mutations, and so forth.

In another embodiment of the invention, natural, modified, or recombinant nucleic acid sequences may be ligated to a heterologous sequence to encode a fusion protein. For example, to screen peptide libraries for inhibitors of polypeptide activity, it may be useful to encode a chimeric protein that can be recognized by a commercially available antibody. A fusion protein may also be engineered to contain a cleavage site located between the polypeptide-encoding sequence and the heterologous protein sequence, so that the polypeptide may be cleaved and purified away from the heterologous moiety.

Sequences encoding a desired polypeptide may be synthesized, in whole or in part, using chemical methods well known in the art (see Caruthers, M. H. et al. (1980) *Nucl. Acids Res. Symp. Ser.* 215-223, Horn, T. et al. (1980) *Nucl. Acids Res. Symp. Ser.* 225-232). Alternatively, the protein itself may be produced using chemical methods to synthesize the amino acid sequence of a polypeptide, or a portion thereof. For example, peptide synthesis can be performed using various solid-phase techniques (Roberge, J. Y. et al. (1995) *Science 269*:202-204) and automated synthesis may be achieved, for example, using the ABI 431A Peptide Synthesizer (Perkin Elmer, Palo Alto, CA).

A newly synthesized peptide may be substantially purified by preparative high performance liquid chromatography (*e.g.*, Creighton, T. (1983) Proteins, Structures and Molecular Principles, WH Freeman and Co., New York, N.Y.) or other comparable techniques available in the art. The composition of the synthetic peptides may be confirmed by amino acid analysis or sequencing (*e.g.*, the Edman degradation procedure). Additionally, the amino acid sequence of a polypeptide, or any part thereof, may be altered during direct synthesis and/or combined using chemical methods with sequences from other proteins, or any part thereof, to produce a variant polypeptide.

In order to express a desired polypeptide, the nucleotide sequences encoding the polypeptide, or functional equivalents, may be inserted into appropriate expression vector, *i.e.*, a vector which contains the necessary elements for the

transcription and translation of the inserted coding sequence. Methods which are well known to those skilled in the art may be used to construct expression vectors containing sequences encoding a polypeptide of interest and appropriate transcriptional and translational control elements. These methods include *in vitro* recombinant DNA techniques, synthetic techniques, and *in vivo* genetic recombination. Such techniques are described, for example, in Sambrook, J. et al. (1989) Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, Plainview, N.Y., and Ausubel, F. M. et al. (1989) Current Protocols in Molecular Biology, John Wiley & Sons, New York. N.Y.

A variety of expression vector/host systems may be utilized to contain and express polynucleotide sequences. These include, but are not limited to, microorganisms such as bacteria transformed with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with yeast expression vectors; insect cell systems infected with virus expression vectors (*e.g.*, baculovirus); plant cell systems transformed with virus expression vectors (*e.g.*, cauliflower mosaic virus, CaMV; tobacco mosaic virus, TMV) or with bacterial expression vectors (*e.g.*, Ti or pBR322 plasmids); or animal cell systems.

The "control elements" or "regulatory sequences" present in an expression vector are those non-translated regions of the vector--enhancers, promoters, 5' and 3' untranslated regions--which interact with host cellular proteins to carry out transcription and translation. Such elements may vary in their strength and specificity. Depending on the vector system and host utilized, any number of suitable transcription and translation elements, including constitutive and inducible promoters, may be used. For example, when cloning in bacterial systems, inducible promoters such as the hybrid lacZ promoter of the pBLUESCRIPT phagemid (Stratagene, La Jolla, Calif.) or pSPORT1 plasmid (Gibco BRL, Gaithersburg, MD) and the like may be used. In mammalian cell systems, promoters from mammalian genes or from mammalian viruses are generally preferred. If it is necessary to generate a cell line that contains multiple copies of the sequence encoding a polypeptide, vectors based on SV40 or EBV may be advantageously used with an appropriate selectable marker.

In bacterial systems, any of a number of expression vectors may be selected depending upon the use intended for the expressed polypeptide. For example, when large quantities are needed, for example for the induction of antibodies, vectors which direct high level expression of fusion proteins that are readily purified may be used. Such vectors include, but are not limited to, the multifunctional *E. coli* cloning and expression vectors such as pBLUESCRIPT (Stratagene), in which the sequence encoding the polypeptide of interest may be ligated into the vector in frame with sequences for the amino-terminal Met and the subsequent 7 residues of .beta.-galactosidase so that a hybrid protein is produced; pIN vectors (Van Heeke, G. and S. M. Schuster (1989) *J. Biol. Chem. 264*:5503-5509); and the like. pGEX Vectors (Promega, Madison, Wis.) may also be used to express foreign polypeptides as fusion proteins with glutathione S-transferase (GST). In general, such fusion proteins are soluble and can easily be purified from lysed cells by adsorption to glutathione-agarose beads followed by elution in the presence of free glutathione. Proteins made in such systems may be designed to include heparin, thrombin, or factor XA protease cleavage sites so that the cloned polypeptide of interest can be released from the GST moiety at will.

In the yeast, Saccharomyces cerevisiae, a number of vectors containing constitutive or inducible promoters such as alpha factor, alcohol oxidase, and PGH may be used. For reviews, see Ausubel et al. (supra) and Grant et al. (1987) *Methods Enzymol.* 153:516-544.

In cases where plant expression vectors are used, the expression of sequences encoding polypeptides may be driven by any of a number of promoters. For example, viral promoters such as the 35S and 19S promoters of CaMV may be used alone or in combination with the omega leader sequence from TMV (Takamatsu, N. (1987) *EMBO J. 6*:307-311. Alternatively, plant promoters such as the small subunit of RUBISCO or heat shock promoters may be used (Coruzzi, G. et al. (1984) *EMBO J. 3*:1671-1680; Broglie, R. et al. (1984) *Science 224*:838-843; and Winter, J. et al. (1991) *Results Probl. Cell Differ. 17*:85-105). These constructs can be introduced into plant cells by direct DNA transformation or pathogen-mediated transfection. Such techniques are described in a number of generally available reviews (see, for example, Hobbs, S. or

Murry, L. E. in McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York, N.Y.; pp. 191-196).

An insect system may also be used to express a polypeptide of interest. For example, in one such system, Autographa californica nuclear polyhedrosis virus (AcNPV) is used as a vector to express foreign genes in Spodoptera frugiperda cells or in Trichoplusia larvae. The sequences encoding the polypeptide may be cloned into a non-essential region of the virus, such as the polyhedrin gene, and placed under control of the polyhedrin promoter. Successful insertion of the polypeptide-encoding sequence will render the polyhedrin gene inactive and produce recombinant virus lacking coat protein. The recombinant viruses may then be used to infect, for example, S. frugiperda cells or Trichoplusia larvae in which the polypeptide of interest may be expressed (Engelhard, E. K. et al. (1994) *Proc. Natl. Acad. Sci. 91* :3224-3227).

In mammalian host cells, a number of viral-based expression systems are generally available. For example, in cases where an adenovirus is used as an expression vector, sequences encoding a polypeptide of interest may be ligated into an adenovirus transcription/translation complex consisting of the late promoter and tripartite leader sequence. Insertion in a non-essential E1 or E3 region of the viral genome may be used to obtain a viable virus which is capable of expressing the polypeptide in infected host cells (Logan, J. and Shenk, T. (1984) *Proc. Natl. Acad. Sci. 81*:3655-3659). In addition, transcription enhancers, such as the Rous sarcoma virus (RSV) enhancer, may be used to increase expression in mammalian host cells.

Specific initiation signals may also be used to achieve more efficient translation of sequences encoding a polypeptide of interest. Such signals include the ATG initiation codon and adjacent sequences. In cases where sequences encoding the polypeptide, its initiation codon, and upstream sequences are inserted into the appropriate expression vector, no additional transcriptional or translational control signals may be needed. However, in cases where only coding sequence, or a portion thereof, is inserted, exogenous translational control signals including the ATG initiation codon should be provided. Furthermore, the initiation codon should be in the correct reading frame to ensure translation of the entire insert. Exogenous translational elements and initiation codons may be of various origins, both natural and synthetic.

The efficiency of expression may be enhanced by the inclusion of enhancers which are appropriate for the particular cell system which is used, such as those described in the literature (Scharf, D. et al. (1994) *Results Probl. Cell Differ. 20*:125-162).

In addition, a host cell strain may be chosen for its ability to modulate the expression of the inserted sequences or to process the expressed protein in the desired fashion. Such modifications of the polypeptide include, but are not limited to, acetylation, carboxylation. glycosylation, phosphorylation, lipidation, and acylation. Post-translational processing which cleaves a "prepro" form of the protein may also be used to facilitate correct insertion, folding and/or function. Different host cells such as CHO, COS, HeLa, MDCK, HEK293, and WI38, which have specific cellular machinery and characteristic mechanisms for such post-translational activities, may be chosen to ensure the correct modification and processing of the foreign protein.

For long-term, high-yield production of recombinant proteins, stable expression is generally preferred. For example, cell lines which stably express a polynucleotide of interest may be transformed using expression vectors which may contain viral origins of replication and/or endogenous expression elements and a selectable marker gene on the same or on a separate vector. Following the introduction of the vector, cells may be allowed to grow for 1-2 days in an enriched media before they are switched to selective media. The purpose of the selectable marker is to confer resistance to selection, and its presence allows growth and recovery of cells which successfully express the introduced sequences. Resistant clones of stably transformed cells may be proliferated using tissue culture techniques appropriate to the cell type.

Any number of selection systems may be used to recover transformed cell lines. These include, but are not limited to, the herpes simplex virus thymidine kinase (Wigler, M. et al. (1977) *Cell 11*:223-32) and adenine phosphoribosyltransferase (Lowy, I. et al. (1990) *Cell 22*:817-23) genes which can be employed in tk.sup.- or aprt.sup.- cells, respectively. Also, antimetabolite, antibiotic or herbicide resistance can be used as the basis for selection; for example, dhfr which confers resistance to methotrexate (Wigler, M. et al. (1980) *Proc. Natl. Acad. Sci. 77*:3567-70); npt, which confers resistance to the aminoglycosides, neomycin and G-418 (Colbere-Garapin, F. et al (1981) *J. Mol. Biol. 150*:1-14); and als or pat, which confer resistance to

chlorsulfuron and phosphinotricin acetyltransferase, respectively (Murry, *supra*). Additional selectable genes have been described, for example, trpB, which allows cells to utilize indole in place of tryptophan, or hisD, which allows cells to utilize histinol in place of histidine (Hartman, S. C. and R. C. Mulligan (1988) *Proc. Natl. Acad. Sci.* *85*:8047-51). The use of visible markers has gained popularity with such markers as anthocyanins, beta-glucuronidase and its substrate GUS, and luciferase and its substrate luciferin, being widely used not only to identify transformants, but also to quantify the amount of transient or stable protein expression attributable to a specific vector system (Rhodes, C. A. et al. (1995) *Methods Mol. Biol. 55*:121-131).

Although the presence/absence of marker gene expression suggests that the gene of interest is also present, its presence and expression may need to be confirmed. For example, if the sequence encoding a polypeptide is inserted within a marker gene sequence, recombinant cells containing sequences can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a polypeptide-encoding sequence under the control of a single promoter. Expression of the marker gene in response to induction or selection usually indicates expression of the tandem gene as well.

Alternatively, host cells that contain and express a desired polynucleotide sequence may be identified by a variety of procedures known to those of skill in the art. These procedures include, but are not limited to, DNA-DNA or DNA-RNA hybridizations and protein bioassay or immunoassay techniques which include, for example, membrane, solution, or chip based technologies for the detection and/or quantification of nucleic acid or protein.

A variety of protocols for detecting and measuring the expression of polynucleotide-encoded products, using either polyclonal or monoclonal antibodies specific for the product are known in the art. Examples include enzyme-linked immunosorbent assay (ELISA), radioimmunoassay (RIA), and fluorescence activated cell sorting (FACS). A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes on a given polypeptide may be preferred for some applications, but a competitive binding assay may also be employed. These and other assays are described, among other places, in Hampton, R. et al. (1990;

Serological Methods, a Laboratory Manual, APS Press, St Paul. Minn.) and Maddox, D. E. et al. (1983; *J. Exp. Med. 158*:1211-1216).

A wide variety of labels and conjugation techniques are known by those skilled in the art and may be used in various nucleic acid and amino acid assays. Means for producing labeled hybridization or PCR probes for detecting sequences related to polynucleotides include oligolabeling, nick translation, end-labeling or PCR amplification using a labeled nucleotide. Alternatively, the sequences, or any portions thereof may be cloned into a vector for the production of an mRNA probe. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes in vitro by addition of an appropriate RNA polymerase such as T7, T3, or SP6 and labeled nucleotides. These procedures may be conducted using a variety of commercially available kits. Suitable reporter molecules or labels, which may be used include radionuclides, enzymes, fluorescent, chemiluminescent, or chromogenic agents as well as substrates, cofactors, inhibitors, magnetic particles, and the like.

Host cells transformed with a polynucleotide sequence of interest may be cultured under conditions suitable for the expression and recovery of the protein from cell culture. The protein produced by a recombinant cell may be secreted or contained intracellularly depending on the sequence and/or the vector used. As will be understood by those of skill in the art, expression vectors containing polynucleotides of the invention may be designed to contain signal sequences which direct secretion of the encoded polypeptide through a prokaryotic or eukaryotic cell membrane. Other recombinant constructions may be used to join sequences encoding a polypeptide of interest to nucleotide sequence encoding a polypeptide domain which will facilitate purification of soluble proteins. Such purification facilitating domains include, but are not limited to, metal chelating peptides such as histidine-tryptophan modules that allow purification on immobilized metals, protein A domains that allow purification on immobilized immunoglobulin, and the domain utilized in the FLAGS extension/affinity purification system (Immunex Corp., Seattle, Wash.). The inclusion of cleavable linker sequences such as those specific for Factor XA or enterokinase (Invitrogen. San Diego, Calif.) between the purification domain and the encoded polypeptide may be used to facilitate purification. One such expression vector provides for expression of a fusion

protein containing a polypeptide of interest and a nucleic acid encoding 6 histidine residues preceding a thioredoxin or an enterokinase cleavage site. The histidine residues facilitate purification on IMIAC (immobilized metal ion affinity chromatography) as described in Porath, J. et al. (1992, *Prot. Exp. Purif. 3*:263-281) while the enterokinase cleavage site provides a means for purifying the desired polypeptide from the fusion protein. A discussion of vectors which contain fusion proteins is provided in Kroll, D. J. et al. (1993; *DNA Cell Biol. 12*:441-453).

In addition to recombinant production methods, polypeptides of the invention, and fragments thereof, may be produced by direct peptide synthesis using solid-phase techniques (Merrifield J. (1963) *J. Am. Chem. Soc. 85*:2149-2154). Protein synthesis may be performed using manual techniques or by automation. Automated synthesis may be achieved, for example, using Applied Biosystems 431A Peptide Synthesizer (Perkin Elmer). Alternatively, various fragments may be chemically synthesized separately and combined using chemical methods to produce the full length molecule.

ANTIBODY COMPOSITIONS, FRAGMENTS THEREOF AND OTHER BINDING AGENTS

According to another aspect, the present invention further provides binding agents, such as antibodies and antigen-binding fragments thereof, that exhibit immunological binding to a tumor polypeptide disclosed herein, or to a portion, variant or derivative thereof. An antibody, or antigen-binding fragment thereof, is said to "specifically bind," "immunogically bind," and/or is "immunologically reactive" to a polypeptide of the invention if it reacts at a detectable level (within, for example, an ELISA assay) with the polypeptide, and does not react detectably with unrelated polypeptides under similar conditions.

Immunological binding, as used in this context, generally refers to the non-covalent interactions of the type which occur between an immunoglobulin molecule and an antigen for which the immunoglobulin is specific. The strength, or affinity of immunological binding interactions can be expressed in terms of the dissociation constant (K_d) of the interaction, wherein a smaller K_d represents a greater affinity. Immunological binding properties of selected polypeptides can be quantified

using methods well known in the art. One such method entails measuring the rates of antigen-binding site/antigen complex formation and dissociation, wherein those rates depend on the concentrations of the complex partners, the affinity of the interaction, and on geometric parameters that equally influence the rate in both directions. Thus, both the "on rate constant" (K_{on}) and the "off rate constant" (K_{off}) can be determined by calculation of the concentrations and the actual rates of association and dissociation. The ratio of K_{off}/K_{on} enables cancellation of all parameters not related to affinity, and is thus equal to the dissociation constant K_d. See, generally, Davies et al. (1990) Annual Rev. Biochem. 59:439-473.

An "antigen-binding site," or "binding portion" of an antibody refers to the part of the immunoglobulin molecule that participates in antigen binding. The antigen binding site is formed by amino acid residues of the N-terminal variable ("V") regions of the heavy ("H") and light ("L") chains. Three highly divergent stretches within the V regions of the heavy and light chains are referred to as "hypervariable regions" which are interposed between more conserved flanking stretches known as "framework regions," or "FRs". Thus the term "FR" refers to amino acid sequences which are naturally found between and adjacent to hypervariable regions in immunoglobulins. In an antibody molecule, the three hypervariable regions of a light chain and the three hypervariable regions of a heavy chain are disposed relative to each other in three dimensional space to form an antigen-binding surface. The antigen-binding surface is complementary to the three-dimensional surface of a bound antigen, and the three hypervariable regions of each of the heavy and light chains are referred to as "complementarity-determining regions," or "CDRs."

Binding agents may be further capable of differentiating between patients with and without a cancer, such as colon cancer, using the representative assays provided herein. For example, antibodies or other binding agents that bind to a tumor protein will preferably generate a signal indicating the presence of a cancer in at least about 20% of patients with the disease, more preferably at least about 30% of patients. Alternatively, or in addition, the antibody will generate a negative signal indicating the absence of the disease in at least about 90% of individuals without the cancer. To determine whether a binding agent satisfies this requirement, biological samples (*e.g.*,

blood, sera, sputum, urine and/or tumor biopsies) from patients with and without a cancer (as determined using standard clinical tests) may be assayed as described herein for the presence of polypeptides that bind to the binding agent. Preferably, a statistically significant number of samples with and without the disease will be assayed. Each binding agent should satisfy the above criteria; however, those of ordinary skill in the art will recognize that binding agents may be used in combination to improve sensitivity.

Any agent that satisfies the above requirements may be a binding agent. For example, a binding agent may be a ribosome, with or without a peptide component, an RNA molecule or a polypeptide. In a preferred embodiment, a binding agent is an antibody or an antigen-binding fragment thereof. Antibodies may be prepared by any of a variety of techniques known to those of ordinary skill in the art. *See, e.g.,* Harlow and Lane, *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory, 1988. In general, antibodies can be produced by cell culture techniques, including the generation of monoclonal antibodies as described herein, or via transfection of antibody genes into suitable bacterial or mammalian cell hosts, in order to allow for the production of recombinant antibodies. In one technique, an immunogen comprising the polypeptide is initially injected into any of a wide variety of mammals (*e.g.,* mice, rats, rabbits, sheep or goats). In this step, the polypeptides of this invention may serve as the immunogen without modification. Alternatively, particularly for relatively short polypeptides, a superior immune response may be elicited if the polypeptide is joined to a carrier protein, such as bovine serum albumin or keyhole limpet hemocyanin. The immunogen is injected into the animal host, preferably according to a predetermined schedule incorporating one or more booster immunizations, and the animals are bled periodically. Polyclonal antibodies specific for the polypeptide may then be purified from such antisera by, for example, affinity chromatography using the polypeptide coupled to a suitable solid support.

Monoclonal antibodies specific for an antigenic polypeptide of interest may be prepared, for example, using the technique of Kohler and Milstein, *Eur. J. Immunol. 6*:511-519, 1976, and improvements thereto. Briefly, these methods involve the preparation of immortal cell lines capable of producing antibodies having the

desired specificity (*i.e.*, reactivity with the polypeptide of interest). Such cell lines may be produced, for example, from spleen cells obtained from an animal immunized as described above. The spleen cells are then immortalized by, for example, fusion with a myeloma cell fusion partner, preferably one that is syngeneic with the immunized animal. A variety of fusion techniques may be employed. For example, the spleen cells and myeloma cells may be combined with a nonionic detergent for a few minutes and then plated at low density on a selective medium that supports the growth of hybrid cells, but not myeloma cells. A preferred selection technique uses HAT (hypoxanthine, aminopterin, thymidine) selection. After a sufficient time, usually about 1 to 2 weeks, colonies of hybrids are observed. Single colonies are selected and their culture supernatants tested for binding activity against the polypeptide. Hybridomas having high reactivity and specificity are preferred.

Monoclonal antibodies may be isolated from the supernatants of growing hybridoma colonies. In addition, various techniques may be employed to enhance the yield, such as injection of the hybridoma cell line into the peritoneal cavity of a suitable vertebrate host, such as a mouse. Monoclonal antibodies may then be harvested from the ascites fluid or the blood. Contaminants may be removed from the antibodies by conventional techniques, such as chromatography, gel filtration, precipitation, and extraction. The polypeptides of this invention may be used in the purification process in, for example, an affinity chromatography step.

A number of therapeutically useful molecules are known in the art which comprise antigen-binding sites that are capable of exhibiting immunological binding properties of an antibody molecule. The proteolytic enzyme papain preferentially cleaves IgG molecules to yield several fragments, two of which (the "F(ab)" fragments) each comprise a covalent heterodimer that includes an intact antigen-binding site. The enzyme pepsin is able to cleave IgG molecules to provide several fragments, including the "$F(ab')_2$ " fragment which comprises both antigen-binding sites. An "Fv" fragment can be produced by preferential proteolytic cleavage of an IgM, and on rare occasions IgG or IgA immunoglobulin molecule. Fv fragments are, however, more commonly derived using recombinant techniques known in the art. The Fv fragment includes a non-covalent $V_H::V_L$ heterodimer including an antigen-binding site which retains much

of the antigen recognition and binding capabilities of the native antibody molecule. Inbar et al. (1972) Proc. Nat. Acad. Sci. USA 69:2659-2662; Hochman et al. (1976) Biochem 15:2706-2710; and Ehrlich et al. (1980) Biochem 19:4091-4096.

A single chain Fv ("sFv") polypeptide is a covalently linked V_H::V_L heterodimer which is expressed from a gene fusion including V_H- and V_L-encoding genes linked by a peptide-encoding linker. Huston et al. (1988) Proc. Nat. Acad. Sci. USA 85(16):5879-5883. A number of methods have been described to discern chemical structures for converting the naturally aggregated--but chemically separated--light and heavy polypeptide chains from an antibody V region into an sFv molecule which will fold into a three dimensional structure substantially similar to the structure of an antigen-binding site. See, *e.g.*, U.S. Pat. Nos. 5,091,513 and 5,132,405, to Huston et al.; and U.S. Pat. No. 4,946,778, to Ladner et al.

Each of the above-described molecules includes a heavy chain and a light chain CDR set, respectively interposed between a heavy chain and a light chain FR set which provide support to the CDRS and define the spatial relationship of the CDRs relative to each other. As used herein, the term "CDR set" refers to the three hypervariable regions of a heavy or light chain V region. Proceeding from the N-terminus of a heavy or light chain, these regions are denoted as "CDR1," "CDR2," and "CDR3" respectively. An antigen-binding site, therefore, includes six CDRs, comprising the CDR set from each of a heavy and a light chain V region. A polypeptide comprising a single CDR, (*e.g.*, a CDR1, CDR2 or CDR3) is referred to herein as a "molecular recognition unit." Crystallographic analysis of a number of antigen-antibody complexes has demonstrated that the amino acid residues of CDRs form extensive contact with bound antigen, wherein the most extensive antigen contact is with the heavy chain CDR3. Thus, the molecular recognition units are primarily responsible for the specificity of an antigen-binding site.

As used herein, the term "FR set" refers to the four flanking amino acid sequences which frame the CDRs of a CDR set of a heavy or light chain V region. Some FR residues may contact bound antigen; however, FRs are primarily responsible for folding the V region into the antigen-binding site, particularly the FR residues directly adjacent to the CDRS. Within FRs, certain amino residues and certain structural

features are very highly conserved. In this regard, all V region sequences contain an internal disulfide loop of around 90 amino acid residues. When the V regions fold into a binding-site, the CDRs are displayed as projecting loop motifs which form an antigen-binding surface. It is generally recognized that there are conserved structural regions of FRs which influence the folded shape of the CDR loops into certain "canonical" structures--regardless of the precise CDR amino acid sequence. Further, certain FR residues are known to participate in non-covalent interdomain contacts which stabilize the interaction of the antibody heavy and light chains.

A number of "humanized" antibody molecules comprising an antigen-binding site derived from a non-human immunoglobulin have been described, including chimeric antibodies having rodent V regions and their associated CDRs fused to human constant domains (Winter et al. (1991) Nature 349:293-299; Lobuglio et al. (1989) Proc. Nat. Acad. Sci. USA 86:4220-4224; Shaw et al. (1987) J Immunol. 138:4534-4538; and Brown et al. (1987) Cancer Res. 47:3577-3583), rodent CDRs grafted into a human supporting FR prior to fusion with an appropriate human antibody constant domain (Riechmann et al. (1988) Nature 332:323-327; Verhoeyen et al. (1988) Science 239:1534-1536; and Jones et al. (1986) Nature 321:522-525), and rodent CDRs supported by recombinantly veneered rodent FRs (European Patent Publication No. 519,596, published Dec. 23, 1992). These "humanized" molecules are designed to minimize unwanted immunological response toward rodent antihuman antibody molecules which limits the duration and effectiveness of therapeutic applications of those moieties in human recipients.

As used herein, the terms "veneered FRs" and "recombinantly veneered FRs" refer to the selective replacement of FR residues from, *e.g.*, a rodent heavy or light chain V region, with human FR residues in order to provide a xenogeneic molecule comprising an antigen-binding site which retains substantially all of the native FR polypeptide folding structure. Veneering techniques are based on the understanding that the ligand binding characteristics of an antigen-binding site are determined primarily by the structure and relative disposition of the heavy and light chain CDR sets within the antigen-binding surface. Davies et al. (1990) Ann. Rev. Biochem. 59:439-473. Thus, antigen binding specificity can be preserved in a humanized antibody only wherein the

CDR structures, their interaction with each other, and their interaction with the rest of the V region domains are carefully maintained. By using veneering techniques, exterior (*e.g.*, solvent-accessible) FR residues which are readily encountered by the immune system are selectively replaced with human residues to provide a hybrid molecule that comprises either a weakly immunogenic, or substantially non-immunogenic veneered surface.

The process of veneering makes use of the available sequence data for human antibody variable domains compiled by Kabat et al., in Sequences of Proteins of Immunological Interest, 4th ed., (U.S. Dept. of Health and Human Services, U.S. Government Printing Office, 1987), updates to the Kabat database, and other accessible U.S. and foreign databases (both nucleic acid and protein). Solvent accessibilities of V region amino acids can be deduced from the known three-dimensional structure for human and murine antibody fragments. There are two general steps in veneering a murine antigen-binding site. Initially, the FRs of the variable domains of an antibody molecule of interest are compared with corresponding FR sequences of human variable domains obtained from the above-identified sources. The most homologous human V regions are then compared residue by residue to corresponding murine amino acids. The residues in the murine FR which differ from the human counterpart are replaced by the residues present in the human moiety using recombinant techniques well known in the art. Residue switching is only carried out with moieties which are at least partially exposed (solvent accessible), and care is exercised in the replacement of amino acid residues which may have a significant effect on the tertiary structure of V region domains, such as proline, glycine and charged amino acids.

In this manner, the resultant "veneered" murine antigen-binding sites are thus designed to retain the murine CDR residues, the residues substantially adjacent to the CDRs, the residues identified as buried or mostly buried (solvent inaccessible), the residues believed to participate in non-covalent (*e.g.*, electrostatic and hydrophobic) contacts between heavy and light chain domains, and the residues from conserved structural regions of the FRs which are believed to influence the "canonical" tertiary structures of the CDR loops. These design criteria are then used to prepare recombinant nucleotide sequences which combine the CDRs of both the heavy and light chain of a

murine antigen-binding site into human-appearing FRs that can be used to transfect mammalian cells for the expression of recombinant human antibodies which exhibit the antigen specificity of the murine antibody molecule.

In another embodiment of the invention, monoclonal antibodies of the present invention may be coupled to one or more therapeutic agents. Suitable agents in this regard include radionuclides, differentiation inducers, drugs, toxins, and derivatives thereof. Preferred radionuclides include ^{90}Y, ^{123}I, ^{125}I, ^{131}I, ^{186}Re, ^{188}Re, ^{211}At, and ^{212}Bi. Preferred drugs include methotrexate, and pyrimidine and purine analogs. Preferred differentiation inducers include phorbol esters and butyric acid. Preferred toxins include ricin, abrin, diptheria toxin, cholera toxin, gelonin, Pseudomonas exotoxin, Shigella toxin, and pokeweed antiviral protein.

A therapeutic agent may be coupled (*e.g.*, covalently bonded) to a suitable monoclonal antibody either directly or indirectly (*e.g.*, via a linker group). A direct reaction between an agent and an antibody is possible when each possesses a substituent capable of reacting with the other. For example, a nucleophilic group, such as an amino or sulfhydryl group, on one may be capable of reacting with a carbonyl-containing group, such as an anhydride or an acid halide, or with an alkyl group containing a good leaving group (*e.g.*, a halide) on the other.

Alternatively, it may be desirable to couple a therapeutic agent and an antibody via a linker group. A linker group can function as a spacer to distance an antibody from an agent in order to avoid interference with binding capabilities. A linker group can also serve to increase the chemical reactivity of a substituent on an agent or an antibody, and thus increase the coupling efficiency. An increase in chemical reactivity may also facilitate the use of agents, or functional groups on agents, which otherwise would not be possible.

It will be evident to those skilled in the art that a variety of bifunctional or polyfunctional reagents, both homo- and hetero-functional (such as those described in the catalog of the Pierce Chemical Co., Rockford, IL), may be employed as the linker group. Coupling may be effected, for example, through amino groups, carboxyl groups, sulfhydryl groups or oxidized carbohydrate residues. There are numerous references describing such methodology, *e.g.*, U.S. Patent No. 4,671,958, to Rodwell et al.

Where a therapeutic agent is more potent when free from the antibody portion of the immunoconjugates of the present invention, it may be desirable to use a linker group which is cleavable during or upon internalization into a cell. A number of different cleavable linker groups have been described. The mechanisms for the intracellular release of an agent from these linker groups include cleavage by reduction of a disulfide bond (*e.g.*, U.S. Patent No. 4,489,710, to Spitler), by irradiation of a photolabile bond (*e.g.*, U.S. Patent No. 4,625,014, to Senter et al.), by hydrolysis of derivatized amino acid side chains (*e.g.*, U.S. Patent No. 4,638,045, to Kohn et al.), by serum complement-mediated hydrolysis (*e.g.*, U.S. Patent No. 4,671,958, to Rodwell et al.), and acid-catalyzed hydrolysis (*e.g.*, U.S. Patent No. 4,569,789, to Blattler et al.).

It may be desirable to couple more than one agent to an antibody. In one embodiment, multiple molecules of an agent are coupled to one antibody molecule. In another embodiment, more than one type of agent may be coupled to one antibody. Regardless of the particular embodiment, immunoconjugates with more than one agent may be prepared in a variety of ways. For example, more than one agent may be coupled directly to an antibody molecule, or linkers that provide multiple sites for attachment can be used. Alternatively, a carrier can be used.

A carrier may bear the agents in a variety of ways, including covalent bonding either directly or via a linker group. Suitable carriers include proteins such as albumins (*e.g.*, U.S. Patent No. 4,507,234, to Kato et al.), peptides and polysaccharides such as aminodextran (*e.g.*, U.S. Patent No. 4,699,784, to Shih et al.). A carrier may also bear an agent by noncovalent bonding or by encapsulation, such as within a liposome vesicle (*e.g.*, U.S. Patent Nos. 4,429,008 and 4,873,088). Carriers specific for radionuclide agents include radiohalogenated small molecules and chelating compounds. For example, U.S. Patent No. 4,735,792 discloses representative radiohalogenated small molecules and their synthesis. A radionuclide chelate may be formed from chelating compounds that include those containing nitrogen and sulfur atoms as the donor atoms for binding the metal, or metal oxide, radionuclide. For example, U.S. Patent No. 4,673,562, to Davison et al. discloses representative chelating compounds and their synthesis.

T CELL COMPOSITIONS

The present invention, in another aspect, provides T cells specific for a tumor polypeptide disclosed herein, or for a variant or derivative thereof. Such cells may generally be prepared *in vitro* or *ex vivo*, using standard procedures. For example, T cells may be isolated from bone marrow, peripheral blood, or a fraction of bone marrow or peripheral blood of a patient, using a commercially available cell separation system, such as the Isolex™ System, available from Nexell Therapeutics, Inc. (Irvine, CA; see also U.S. Patent No. 5,240,856; U.S. Patent No. 5,215,926; WO 89/06280; WO 91/16116 and WO 92/07243). Alternatively, T cells may be derived from related or unrelated humans, non-human mammals, cell lines or cultures.

T cells may be stimulated with a polypeptide, polynucleotide encoding a polypeptide and/or an antigen presenting cell (APC) that expresses such a polypeptide. Such stimulation is performed under conditions and for a time sufficient to permit the generation of T cells that are specific for the polypeptide of interest. Preferably, a tumor polypeptide or polynucleotide of the invention is present within a delivery vehicle, such as a microsphere, to facilitate the generation of specific T cells.

T cells are considered to be specific for a polypeptide of the present invention if the T cells specifically proliferate, secrete cytokines or kill target cells coated with the polypeptide or expressing a gene encoding the polypeptide. T cell specificity may be evaluated using any of a variety of standard techniques. For example, within a chromium release assay or proliferation assay, a stimulation index of more than two fold increase in lysis and/or proliferation, compared to negative controls, indicates T cell specificity. Such assays may be performed, for example, as described in Chen et al., *Cancer Res.* *54*:1065-1070, 1994. Alternatively, detection of the proliferation of T cells may be accomplished by a variety of known techniques. For example, T cell proliferation can be detected by measuring an increased rate of DNA synthesis (*e.g.*, by pulse-labeling cultures of T cells with tritiated thymidine and measuring the amount of tritiated thymidine incorporated into DNA). Contact with a tumor polypeptide (100 ng/ml - 100 μg/ml, preferably 200 ng/ml - 25 μg/ml) for 3 - 7 days will typically result in at least a two fold increase in proliferation of the T cells. Contact as described above for 2-3 hours should result in activation of the T cells, as

measured using standard cytokine assays in which a two fold increase in the level of cytokine release (*e.g.*, TNF or IFN-γ) is indicative of T cell activation (*see* Coligan et al., Current Protocols in Immunology, vol. 1, Wiley Interscience (Greene 1998)). T cells that have been activated in response to a tumor polypeptide, polynucleotide or polypeptide-expressing APC may be $CD4^+$ and/or $CD8^+$. Tumor polypeptide-specific T cells may be expanded using standard techniques. Within preferred embodiments, the T cells are derived from a patient, a related donor or an unrelated donor, and are administered to the patient following stimulation and expansion.

For therapeutic purposes, $CD4^+$ or $CD8^+$ T cells that proliferate in response to a tumor polypeptide, polynucleotide or APC can be expanded in number either *in vitro* or *in vivo*. Proliferation of such T cells *in vitro* may be accomplished in a variety of ways. For example, the T cells can be re-exposed to a tumor polypeptide, or a short peptide corresponding to an immunogenic portion of such a polypeptide, with or without the addition of T cell growth factors, such as interleukin-2, and/or stimulator cells that synthesize a tumor polypeptide. Alternatively, one or more T cells that proliferate in the presence of the tumor polypeptide can be expanded in number by cloning. Methods for cloning cells are well known in the art, and include limiting dilution.

T CELL RECEPTOR COMPOSITIONS

The T cell receptor (TCR) consists of 2 different, highly variable polypeptide chains, termed the T-cell receptor α and β chains, that are linked by a disulfide bond (Janeway, Travers, Walport. *Immunobiology.* Fourth Ed., 148-159. Elsevier Science Ltd/Garland Publishing. 1999). The α/β heterodimer complexes with the invariant CD3 chains at the cell membrane. This complex recognizes specific antigenic peptides bound to MHC molecules. The enormous diversity of TCR specificities is generated much like immunoglobulin diversity, through somatic gene rearrangement. The β chain genes contain over 50 variable (V), 2 diversity (D), over 10 joining (J) segments, and 2 constant region segments (C). The α chain genes contain over 70 V segments, and over 60 J segments but no D segments, as well as one C segment. During T cell development in the thymus, the D to J gene rearrangement of

the β chain occurs, followed by the V gene segment rearrangement to the DJ. This functional VDJ_β exon is transcribed and spliced to join to a C_β. For the α chain, a V_α gene segment rearranges to a J_α gene segment to create the functional exon that is then transcribed and spliced to the C_α. Diversity is further increased during the recombination process by the random addition of P and N-nucleotides between the V, D, and J segments of the β chain and between the V and J segments in the α chain (Janeway, Travers, Walport. *Immunobiology.* Fourth Ed., 98 and 150. Elsevier Science Ltd/Garland Publishing. 1999).

The present invention, in another aspect, provides TCRs specific for a polypeptide disclosed herein, or for a variant or derivative thereof. In accordance with the present invention, polynucleotide and amino acid sequences are provided for the V-J or V-D-J junctional regions or parts thereof for the alpha and beta chains of the T-cell receptor which recognize tumor polypeptides described herein. In general, this aspect of the invention relates to T-cell receptors which recognize or bind tumor polypeptides presented in the context of MHC. In a preferred embodiment the tumor antigens recognized by the T-cell receptors comprise a polypeptide of the present invention. For example, cDNA encoding a TCR specific for a _tumor peptide can be isolated from T cells specific for a tumor polypeptide using standard molecular biological and recombinant DNA techniques.

This invention further includes the T-cell receptors or analogs thereof having substantially the same function or activity as the T-cell receptors of this invention which recognize or bind tumor polypeptides. Such receptors include, but are not limited to, a fragment of the receptor, or a substitution, addition or deletion mutant of a T-cell receptor provided herein. This invention also encompasses polypeptides or peptides that are substantially homologous to the T-cell receptors provided herein or that retain substantially the same activity. The term "analog" includes any protein or polypeptide having an amino acid residue sequence substantially identical to the T-cell receptors provided herein in which one or more residues, preferably no more than 5 residues, more preferably no more than 25 residues have been conservatively substituted with a functionally similar residue and which displays the functional aspects of the T-cell receptor as described herein.

The present invention further provides for suitable mammalian host cells, for example, non-specific T cells, that are transfected with a polynucleotide encoding TCRs specific for a polypeptide described herein, thereby rendering the host cell specific for the polypeptide. The α and β chains of the TCR may be contained on separate expression vectors or alternatively, on a single expression vector that also contains an internal ribosome entry site (IRES) for cap-independent translation of the gene downstream of the IRES. Said host cells expressing TCRs specific for the polypeptide may be used, for example, for adoptive immunotherapy of colon cancer as discussed further below.

In further aspects of the present invention, cloned TCRs specific for a polypeptide recited herein may be used in a kit for the diagnosis of colon cancer. For example, the nucleic acid sequence or portions thereof, of tumor-specific TCRs can be used as probes or primers for the detection of expression of the rearranged genes encoding the specific TCR in a biological sample. Therefore, the present invention further provides for an assay for detecting messenger RNA or DNA encoding the TCR specific for a polypeptide.

PHARMACEUTICAL COMPOSITIONS

In additional embodiments, the present invention concerns formulation of one or more of the polynucleotide, polypeptide, T-cell, TCR, and/or antibody compositions disclosed herein in pharmaceutically-acceptable carriers for administration to a cell or an animal, either alone, or in combination with one or more other modalities of therapy.

It will be understood that, if desired, a composition as disclosed herein may be administered in combination with other agents as well, such as, *e.g.*, other proteins or polypeptides or various pharmaceutically-active agents. In fact, there is virtually no limit to other components that may also be included, given that the additional agents do not cause a significant adverse effect upon contact with the target cells or host tissues. The compositions may thus be delivered along with various other agents as required in the particular instance. Such compositions may be purified from host cells or other biological sources, or alternatively may be chemically synthesized as

described herein. Likewise, such compositions may further comprise substituted or derivatized RNA or DNA compositions.

Therefore, in another aspect of the present invention, pharmaceutical compositions are provided comprising one or more of the polynucleotide, polypeptide, antibody, TCR, and/or T-cell compositions described herein in combination with a physiologically acceptable carrier. In certain preferred embodiments, the pharmaceutical compositions of the invention comprise immunogenic polynucleotide and/or polypeptide compositions of the invention for use in prophylactic and theraputic vaccine applications. Vaccine preparation is generally described in, for example, M.F. Powell and M.J. Newman, eds., "Vaccine Design (the subunit and adjuvant approach)," Plenum Press (NY, 1995). Generally, such compositions will comprise one or more polynucleotide and/or polypeptide compositions of the present invention in combination with one or more immunostimulants.

It will be apparent that any of the pharmaceutical compositions described herein can contain pharmaceutically acceptable salts of the polynucleotides and polypeptides of the invention. Such salts can be prepared, for example, from pharmaceutically acceptable non-toxic bases, including organic bases (*e.g.,* salts of primary, secondary and tertiary amines and basic amino acids) and inorganic bases (*e.g.,* sodium, potassium, lithium, ammonium, calcium and magnesium salts).

In another embodiment, illustrative immunogenic compositions, *e.g.*, vaccine compositions, of the present invention comprise DNA encoding one or more of the polypeptides as described above, such that the polypeptide is generated *in situ*. As noted above, the polynucleotide may be administered within any of a variety of delivery systems known to those of ordinary skill in the art. Indeed, numerous gene delivery techniques are well known in the art, such as those described by Rolland, *Crit. Rev. Therap. Drug Carrier Systems 15*:143-198, 1998, and references cited therein. Appropriate polynucleotide expression systems will, of course, contain the necessary regulatory DNA regulatory sequences for expression in a patient (such as a suitable promoter and terminating signal). Alternatively, bacterial delivery systems may involve the administration of a bacterium (such as *Bacillus-Calmette-Guerrin*) that expresses an immunogenic portion of the polypeptide on its cell surface or secretes such an epitope.

Therefore, in certain embodiments, polynucleotides encoding immunogenic polypeptides described herein are introduced into suitable mammalian host cells for expression using any of a number of known viral-based systems. In one illustrative embodiment, retroviruses provide a convenient and effective platform for gene delivery systems. A selected nucleotide sequence encoding a polypeptide of the present invention can be inserted into a vector and packaged in retroviral particles using techniques known in the art. The recombinant virus can then be isolated and delivered to a subject. A number of illustrative retroviral systems have been described (*e.g.*, U.S. Pat. No. 5,219,740; Miller and Rosman (1989) BioTechniques 7:980-990; Miller, A. D. (1990) Human Gene Therapy 1:5-14; Scarpa et al. (1991) Virology 180:849-852; Burns et al. (1993) Proc. Natl. Acad. Sci. USA 90:8033-8037; and Boris-Lawrie and Temin (1993) Cur. Opin. Genet. Develop. 3:102-109.

In addition, a number of illustrative adenovirus-based systems have also been described. Unlike retroviruses which integrate into the host genome, adenoviruses persist extrachromosomally thus minimizing the risks associated with insertional mutagenesis (Haj-Ahmad and Graham (1986) J. Virol. 57:267-274; Bett et al. (1993) J. Virol. 67:5911-5921; Mittereder et al. (1994) Human Gene Therapy 5:717-729; Seth et al. (1994) J. Virol. 68:933-940; Barr et al. (1994) Gene Therapy 1:51-58; Berkner, K. L. (1988) BioTechniques 6:616-629; and Rich et al. (1993) Human Gene Therapy 4:461-476).

Various adeno-associated virus (AAV) vector systems have also been developed for polynucleotide delivery. AAV vectors can be readily constructed using techniques well known in the art. See, *e.g.*, U.S. Pat. Nos. 5,173,414 and 5,139,941; International Publication Nos. WO 92/01070 and WO 93/03769; Lebkowski et al. (1988) Molec. Cell. Biol. 8:3988-3996; Vincent et al. (1990) Vaccines 90 (Cold Spring Harbor Laboratory Press); Carter, B. J. (1992) Current Opinion in Biotechnology 3:533-539; Muzyczka, N. (1992) Current Topics in Microbiol. and Immunol. 158:97-129; Kotin, R. M. (1994) Human Gene Therapy 5:793-801; Shelling and Smith (1994) Gene Therapy 1:165-169; and Zhou et al. (1994) J. Exp. Med. 179:1867-1875.

Additional viral vectors useful for delivering the polynucleotides encoding polypeptides of the present invention by gene transfer include those derived

from the pox family of viruses, such as vaccinia virus and avian poxvirus. By way of example, vaccinia virus recombinants expressing the novel molecules can be constructed as follows. The DNA encoding a polypeptide is first inserted into an appropriate vector so that it is adjacent to a vaccinia promoter and flanking vaccinia DNA sequences, such as the sequence encoding thymidine kinase (TK). This vector is then used to transfect cells which are simultaneously infected with vaccinia. Homologous recombination serves to insert the vaccinia promoter plus the gene encoding the polypeptide of interest into the viral genome. The resulting TK.sup.(-) recombinant can be selected by culturing the cells in the presence of 5-bromodeoxyuridine and picking viral plaques resistant thereto.

A vaccinia-based infection/transfection system can be conveniently used to provide for inducible, transient expression or coexpression of one or more polypeptides described herein in host cells of an organism. In this particular system, cells are first infected in vitro with a vaccinia virus recombinant that encodes the bacteriophage T7 RNA polymerase. This polymerase displays exquisite specificity in that it only transcribes templates bearing T7 promoters. Following infection, cells are transfected with the polynucleotide or polynucleotides of interest, driven by a T7 promoter. The polymerase expressed in the cytoplasm from the vaccinia virus recombinant transcribes the transfected DNA into RNA which is then translated into polypeptide by the host translational machinery. The method provides for high level, transient, cytoplasmic production of large quantities of RNA and its translation products. See, *e.g.*, Elroy-Stein and Moss, Proc. Natl. Acad. Sci. USA (1990) 87:6743-6747; Fuerst et al. Proc. Natl. Acad. Sci. USA (1986) 83:8122-8126.

Alternatively, avipoxviruses, such as the fowlpox and canarypox viruses, can also be used to deliver the coding sequences of interest. Recombinant avipox viruses, expressing immunogens from mammalian pathogens, are known to confer protective immunity when administered to non-avian species. The use of an Avipox vector is particularly desirable in human and other mammalian species since members of the Avipox genus can only productively replicate in susceptible avian species and therefore are not infective in mammalian cells. Methods for producing recombinant Avipoxviruses are known in the art and employ genetic recombination, as described

above with respect to the production of vaccinia viruses. See, *e.g.*, WO 91/12882; WO 89/03429; and WO 92/03545.

Any of a number of alphavirus vectors can also be used for delivery of polynucleotide compositions of the present invention, such as those vectors described in U.S. Patent Nos. 5,843,723; 6,015,686; 6,008,035 and 6,015,694. Certain vectors based on Venezuelan Equine Encephalitis (VEE) can also be used, illustrative examples of which can be found in U.S. Patent Nos. 5,505,947 and 5,643,576.

Moreover, molecular conjugate vectors, such as the adenovirus chimeric vectors described in Michael et al. J. Biol. Chem. (1993) 268:6866-6869 and Wagner et al. Proc. Natl. Acad. Sci. USA (1992) 89:6099-6103, can also be used for gene delivery under the invention.

Additional illustrative information on these and other known viral-based delivery systems can be found, for example, in Fisher-Hoch et al., *Proc. Natl. Acad. Sci. USA 86*:317-321, 1989; Flexner et al., *Ann. N.Y. Acad. Sci. 569*:86-103, 1989; Flexner et al., *Vaccine 8*:17-21, 1990; U.S. Patent Nos. 4,603,112, 4,769,330, and 5,017,487; WO 89/01973; U.S. Patent No. 4,777,127; GB 2,200,651; EP 0,345,242; WO 91/02805; Berkner, *Biotechniques 6*:616-627, 1988; Rosenfeld et al., *Science 252*:431-434, 1991; Kolls et al., *Proc. Natl. Acad. Sci. USA 91*:215-219, 1994; Kass-Eisler et al., *Proc. Natl. Acad. Sci. USA 90*:11498-11502, 1993; Guzman et al., *Circulation 88*:2838-2848, 1993; and Guzman et al., *Cir. Res. 73*:1202-1207, 1993.

In certain embodiments, a polynucleotide may be integrated into the genome of a target cell. This integration may be in the specific location and orientation *via* homologous recombination (gene replacement) or it may be integrated in a random, non-specific location (gene augmentation). In yet further embodiments, the polynucleotide may be stably maintained in the cell as a separate, episomal segment of DNA. Such polynucleotide segments or "episomes" encode sequences sufficient to permit maintenance and replication independent of or in synchronization with the host cell cycle. The manner in which the expression construct is delivered to a cell and where in the cell the polynucleotide remains is dependent on the type of expression construct employed.

In another embodiment of the invention, a polynucleotide is administered/delivered as "naked" DNA, for example as described in Ulmer et al., *Science 259*:1745-1749, 1993 and reviewed by Cohen, *Science 259*:1691-1692, 1993. The uptake of naked DNA may be increased by coating the DNA onto biodegradable beads, which are efficiently transported into the cells.

In still another embodiment, a composition of the present invention can be delivered via a particle bombardment approach, many of which have been described. In one illustrative example, gas-driven particle acceleration can be achieved with devices such as those manufactured by Powderject Pharmaceuticals PLC (Oxford, UK) and Powderject Vaccines Inc. (Madison, WI), some examples of which are described in U.S. Patent Nos. 5,846,796; 6,010,478; 5,865,796; 5,584,807; and EP Patent No. 0500 799. This approach offers a needle-free delivery approach wherein a dry powder formulation of microscopic particles, such as polynucleotide or polypeptide particles, are accelerated to high speed within a helium gas jet generated by a hand held device, propelling the particles into a target tissue of interest.

In a related embodiment, other devices and methods that may be useful for gas-driven needle-less injection of compositions of the present invention include those provided by Bioject, Inc. (Portland, OR), some examples of which are described in U.S. Patent Nos. 4,790,824; 5,064,413; 5,312,335; 5,383,851; 5,399,163; 5,520,639 and 5,993,412.

According to another embodiment, the pharmaceutical compositions described herein will comprise one or more immunostimulants in addition to the immunogenic polynucleotide, polypeptide, antibody, T-cell, TCR, and/or APC compositions of this invention. An immunostimulant refers to essentially any substance that enhances or potentiates an immune response (antibody and/or cell-mediated) to an exogenous antigen. One preferred type of immunostimulant comprises an adjuvant. Many adjuvants contain a substance designed to protect the antigen from rapid catabolism, such as aluminum hydroxide or mineral oil, and a stimulator of immune responses, such as lipid A, *Bortadella pertussis* or *Mycobacterium tuberculosis* derived proteins. Certain adjuvants are commercially available as, for example, Freund's Incomplete Adjuvant and Complete Adjuvant (Difco Laboratories, Detroit, MI); Merck

Adjuvant 65 (Merck and Company, Inc., Rahway, NJ); AS-2 (SmithKline Beecham, Philadelphia, PA); aluminum salts such as aluminum hydroxide gel (alum) or aluminum phosphate; salts of calcium, iron or zinc; an insoluble suspension of acylated tyrosine; acylated sugars; cationically or anionically derivatized polysaccharides; polyphosphazenes; biodegradable microspheres; monophosphoryl lipid A and quil A. Cytokines, such as GM-CSF, interleukin-2, -7, -12, and other like growth factors, may also be used as adjuvants.

Within certain embodiments of the invention, the adjuvant composition is preferably one that induces an immune response predominantly of the Th1 type. High levels of Th1-type cytokines (*e.g.,* IFN-γ, TNFα, IL-2 and IL-12) tend to favor the induction of cell mediated immune responses to an administered antigen. In contrast, high levels of Th2-type cytokines (*e.g.,* IL-4, IL-5, IL-6 and IL-10) tend to favor the induction of humoral immune responses. Following application of a vaccine as provided herein, a patient will support an immune response that includes Th1- and Th2-type responses. Within a preferred embodiment, in which a response is predominantly Th1-type, the level of Th1-type cytokines will increase to a greater extent than the level of Th2-type cytokines. The levels of these cytokines may be readily assessed using standard assays. For a review of the families of cytokines, see Mosmann and Coffman, *Ann. Rev. Immunol. 7*:145-173, 1989.

Certain preferred adjuvants for eliciting a predominantly Th1-type response include, for example, a combination of monophosphoryl lipid A, preferably 3-de-O-acylated monophosphoryl lipid A, together with an aluminum salt. MPL® adjuvants are available from Corixa Corporation (Seattle, WA; *see*, for example, US Patent Nos. 4,436,727; 4,877,611; 4,866,034 and 4,912,094). CpG-containing oligonucleotides (in which the CpG dinucleotide is unmethylated) also induce a predominantly Th1 response. Such oligonucleotides are well known and are described, for example, in WO 96/02555, WO 99/33488 and U.S. Patent Nos. 6,008,200 and 5,856,462. Immunostimulatory DNA sequences are also described, for example, by Sato et al., *Science 273*:352, 1996. Another preferred adjuvant comprises a saponin, such as Quil A, or derivatives thereof, including QS21 and QS7 (Aquila Biopharmaceuticals Inc., Framingham, MA); Escin; Digitonin; or *Gypsophila* or

Chenopodium quinoa saponins . Other preferred formulations include more than one saponin in the adjuvant combinations of the present invention, for example combinations of at least two of the following group comprising QS21, QS7, Quil A, β-escin, or digitonin.

Alternatively the saponin formulations may be combined with vaccine vehicles composed of chitosan or other polycationic polymers, polylactide and polylactide-co-glycolide particles, poly-N-acetyl glucosamine-based polymer matrix, particles composed of polysaccharides or chemically modified polysaccharides, liposomes and lipid-based particles, particles composed of glycerol monoesters, etc. The saponins may also be formulated in the presence of cholesterol to form particulate structures such as liposomes or ISCOMs. Furthermore, the saponins may be formulated together with a polyoxyethylene ether or ester, in either a non-particulate solution or suspension, or in a particulate structure such as a paucilamelar liposome or ISCOM. The saponins may also be formulated with excipients such as Carbopol^{R} to increase viscosity, or may be formulated in a dry powder form with a powder excipient such as lactose.

In one preferred embodiment, the adjuvant system includes the combination of a monophosphoryl lipid A and a saponin derivative, such as the combination of QS21 and 3D-MPL® adjuvant, as described in WO 94/00153, or a less reactogenic composition where the QS21 is quenched with cholesterol, as described in WO 96/33739. Other preferred formulations comprise an oil-in-water emulsion and tocopherol. Another particularly preferred adjuvant formulation employing QS21, 3D-MPL® adjuvant and tocopherol in an oil-in-water emulsion is described in WO 95/17210.

Another enhanced adjuvant system involves the combination of a CpG-containing oligonucleotide and a saponin derivative particularly the combination of CpG and QS21 is disclosed in WO 00/09159. Preferably the formulation additionally comprises an oil in water emulsion and tocopherol.

Additional illustrative adjuvants for use in the pharmaceutical compositions of the invention include Montanide ISA 720 (Seppic, France), SAF (Chiron, California, United States), ISCOMS (CSL), MF-59 (Chiron), the SBAS series

of adjuvants (*e.g.,* SBAS-2 or SBAS-4, available from SmithKline Beecham, Rixensart, Belgium), Detox (Enhanzyn®) (Corixa, Hamilton, MT), RC-529 (Corixa, Hamilton, MT) and other aminoalkyl glucosaminide 4-phosphates (AGPs), such as those described in pending U.S. Patent Application Serial Nos. 08/853,826 and 09/074,720, the disclosures of which are incorporated herein by reference in their entireties, and polyoxyethylene ether adjuvants such as those described in WO 99/52549A1.

Other preferred adjuvants include adjuvant molecules of the general formula

(I): $HO(CH_2CH_2O)_n$-A-R,

wherein, n is 1-50, A is a bond or –C(O)-, R is C_{1-50} alkyl or Phenyl C_{1-50} alkyl.

One embodiment of the present invention consists of a vaccine formulation comprising a polyoxyethylene ether of general formula (I), wherein *n* is between 1 and 50, preferably 4-24, most preferably 9; the *R* component is C_{1-50}, preferably C_4-C_{20} alkyl and most preferably C_{12} alkyl, and *A* is a bond. The concentration of the polyoxyethylene ethers should be in the range 0.1-20%, preferably from 0.1-10%, and most preferably in the range 0.1-1%. Preferred polyoxyethylene ethers are selected from the following group: polyoxyethylene-9-lauryl ether, polyoxyethylene-9-steoryl ether, polyoxyethylene-8-steoryl ether, polyoxyethylene-4-lauryl ether, polyoxyethylene-35-lauryl ether, and polyoxyethylene-23-lauryl ether. Polyoxyethylene ethers such as polyoxyethylene lauryl ether are described in the Merck index (12^{th} edition: entry 7717). These adjuvant molecules are described in WO 99/52549.

The polyoxyethylene ether according to the general formula (I) above may, if desired, be combined with another adjuvant. For example, a preferred adjuvant combination is preferably with CpG as described in the pending UK patent application GB 9820956.2.

According to another embodiment of this invention, an immunogenic composition described herein is delivered to a host via antigen presenting cells (APCs), such as dendritic cells, macrophages, B cells, monocytes and other cells that may be engineered to be efficient APCs. Such cells may, but need not, be genetically modified to increase the capacity for presenting the antigen, to improve activation and/or

maintenance of the T cell response, to have anti-tumor effects *per se* and/or to be immunologically compatible with the receiver (*i.e.,* matched HLA haplotype). APCs may generally be isolated from any of a variety of biological fluids and organs, including tumor and peritumoral tissues, and may be autologous, allogeneic, syngeneic or xenogeneic cells.

Certain preferred embodiments of the present invention use dendritic cells or progenitors thereof as antigen-presenting cells. Dendritic cells are highly potent APCs (Banchereau and Steinman, *Nature 392*:245-251, 1998) and have been shown to be effective as a physiological adjuvant for eliciting prophylactic or therapeutic antitumor immunity (*see* Timmerman and Levy, *Ann. Rev. Med. 50*:507-529, 1999). In general, dendritic cells may be identified based on their typical shape (stellate *in situ*, with marked cytoplasmic processes (dendrites) visible *in vitro*), their ability to take up, process and present antigens with high efficiency and their ability to activate naïve T cell responses. Dendritic cells may, of course, be engineered to express specific cell-surface receptors or ligands that are not commonly found on dendritic cells *in vivo* or *ex vivo*, and such modified dendritic cells are contemplated by the present invention. As an alternative to dendritic cells, secreted vesicles antigen-loaded dendritic cells (called exosomes) may be used within a vaccine (*see* Zitvogel et al., *Nature Med. 4*:594-600, 1998).

Dendritic cells and progenitors may be obtained from peripheral blood, bone marrow, tumor-infiltrating cells, peritumoral tissues-infiltrating cells, lymph nodes, spleen, skin, umbilical cord blood or any other suitable tissue or fluid. For example, dendritic cells may be differentiated *ex vivo* by adding a combination of cytokines such as GM-CSF, IL-4, IL-13 and/or TNFα to cultures of monocytes harvested from peripheral blood. Alternatively, CD34 positive cells harvested from peripheral blood, umbilical cord blood or bone marrow may be differentiated into dendritic cells by adding to the culture medium combinations of GM-CSF, IL-3, TNFα, CD40 ligand, LPS, flt3 ligand and/or other compound(s) that induce differentiation, maturation and proliferation of dendritic cells.

Dendritic cells are conveniently categorized as "immature" and "mature" cells, which allows a simple way to discriminate between two well characterized

phenotypes. However, this nomenclature should not be construed to exclude all possible intermediate stages of differentiation. Immature dendritic cells are characterized as APC with a high capacity for antigen uptake and processing, which correlates with the high expression of Fcγ receptor and mannose receptor. The mature phenotype is typically characterized by a lower expression of these markers, but a high expression of cell surface molecules responsible for T cell activation such as class I and class II MHC, adhesion molecules (*e.g.,* CD54 and CD11) and costimulatory molecules (*e.g.,* CD40, CD80, CD86 and 4-1BB).

APCs may generally be transfected with a polynucleotide of the invention (or portion or other variant thereof) such that the encoded polypeptide, or an immunogenic portion thereof, is expressed on the cell surface. Such transfection may take place *ex vivo*, and a pharmaceutical composition comprising such transfected cells may then be used for therapeutic purposes, as described herein. Alternatively, a gene delivery vehicle that targets a dendritic or other antigen presenting cell may be administered to a patient, resulting in transfection that occurs *in vivo*. *In vivo* and *ex vivo* transfection of dendritic cells, for example, may generally be performed using any methods known in the art, such as those described in WO 97/24447, or the gene gun approach described by Mahvi et al., *Immunology and cell Biology* 75:456-460, 1997. Antigen loading of dendritic cells may be achieved by incubating dendritic cells or progenitor cells with the tumor polypeptide, DNA (naked or within a plasmid vector) or RNA; or with antigen-expressing recombinant bacterium or viruses (*e.g.,* vaccinia, fowlpox, adenovirus or lentivirus vectors). Prior to loading, the polypeptide may be covalently conjugated to an immunological partner that provides T cell help (*e.g.,* a carrier molecule). Alternatively, a dendritic cell may be pulsed with a non-conjugated immunological partner, separately or in the presence of the polypeptide.

While any suitable carrier known to those of ordinary skill in the art may be employed in the pharmaceutical compositions of this invention, the type of carrier will typically vary depending on the mode of administration. Compositions of the present invention may be formulated for any appropriate manner of administration, including for example, topical, oral, nasal, mucosal, intravenous, intracranial, intraperitoneal, subcutaneous and intramuscular administration.

Carriers for use within such pharmaceutical compositions are biocompatible, and may also be biodegradable. In certain embodiments, the formulation preferably provides a relatively constant level of active component release. In other embodiments, however, a more rapid rate of release immediately upon administration may be desired. The formulation of such compositions is well within the level of ordinary skill in the art using known techniques. Illustrative carriers useful in this regard include microparticles of poly(lactide-co-glycolide), polyacrylate, latex, starch, cellulose, dextran and the like. Other illustrative delayed-release carriers include supramolecular biovectors, which comprise a non-liquid hydrophilic core (*e.g.,* a cross-linked polysaccharide or oligosaccharide) and, optionally, an external layer comprising an amphiphilic compound, such as a phospholipid (*see e.g.,* U.S. Patent No. 5,151,254 and PCT applications WO 94/20078, WO/94/23701 and WO 96/06638). The amount of active compound contained within a sustained release formulation depends upon the site of implantation, the rate and expected duration of release and the nature of the condition to be treated or prevented.

In another illustrative embodiment, biodegradable microspheres (*e.g.,* polylactate polyglycolate) are employed as carriers for the compositions of this invention. Suitable biodegradable microspheres are disclosed, for example, in U.S. Patent Nos. 4,897,268; 5,075,109; 5,928,647; 5,811,128; 5,820,883; 5,853,763; 5,814,344, 5,407,609 and 5,942,252. Modified hepatitis B core protein carrier systems. such as described in WO/99 40934, and references cited therein, will also be useful for many applications. Another illustrative carrier/delivery system employs a carrier comprising particulate-protein complexes, such as those described in U.S. Patent No. 5,928,647, which are capable of inducing a class I-restricted cytotoxic T lymphocyte responses in a host.

In another illustrative embodiment, calcium phosphate core particles are employed as carriers, vaccine adjuvants, or as controlled release matrices for the compositions of this invention. Exemplary calcium phosphate particles are disclosed, for example, in published patent application No. WO/0046147.

The pharmaceutical compositions of the invention will often further comprise one or more buffers (*e.g.,* neutral buffered saline or phosphate buffered

saline), carbohydrates (*e.g.*, glucose, mannose, sucrose or dextrans), mannitol, proteins, polypeptides or amino acids such as glycine, antioxidants, bacteriostats, chelating agents such as EDTA or glutathione, adjuvants (*e.g.*, aluminum hydroxide), solutes that render the formulation isotonic, hypotonic or weakly hypertonic with the blood of a recipient, suspending agents, thickening agents and/or preservatives. Alternatively, compositions of the present invention may be formulated as a lyophilizate.

The pharmaceutical compositions described herein may be presented in unit-dose or multi-dose containers, such as sealed ampoules or vials. Such containers are typically sealed in such a way to preserve the sterility and stability of the formulation until use. In general, formulations may be stored as suspensions, solutions or emulsions in oily or aqueous vehicles. Alternatively, a pharmaceutical composition may be stored in a freeze-dried condition requiring only the addition of a sterile liquid carrier immediately prior to use.

The development of suitable dosing and treatment regimens for using the particular compositions described herein in a variety of treatment regimens, including *e.g.*, oral, parenteral, intravenous, intranasal, and intramuscular administration and formulation, is well known in the art, some of which are briefly discussed below for general purposes of illustration.

In certain applications, the pharmaceutical compositions disclosed herein may be delivered *via* oral administration to an animal. As such, these compositions may be formulated with an inert diluent or with an assimilable edible carrier, or they may be enclosed in hard- or soft-shell gelatin capsule, or they may be compressed into tablets, or they may be incorporated directly with the food of the diet.

The active compounds may even be incorporated with excipients and used in the form of ingestible tablets, buccal tables, troches, capsules, elixirs, suspensions, syrups, wafers, and the like (see, for example, Mathiowitz *et al.*, Nature 1997 Mar 27;386(6623):410-4; Hwang *et al.*, Crit Rev Ther Drug Carrier Syst 1998;15(3):243-84; U. S. Patent 5,641,515; U. S. Patent 5,580,579 and U. S. Patent 5,792,451). Tablets, troches, pills, capsules and the like may also contain any of a variety of additional components, for example, a binder, such as gum tragacanth, acacia, cornstarch, or gelatin; excipients, such as dicalcium phosphate; a disintegrating

agent, such as corn starch, potato starch, alginic acid and the like; a lubricant, such as magnesium stearate; and a sweetening agent, such as sucrose, lactose or saccharin may be added or a flavoring agent, such as peppermint, oil of wintergreen, or cherry flavoring. When the dosage unit form is a capsule, it may contain, in addition to materials of the above type, a liquid carrier. Various other materials may be present as coatings or to otherwise modify the physical form of the dosage unit. For instance, tablets, pills, or capsules may be coated with shellac, sugar, or both. Of course, any material used in preparing any dosage unit form should be pharmaceutically pure and substantially non-toxic in the amounts employed. In addition, the active compounds may be incorporated into sustained-release preparation and formulations.

Typically, these formulations will contain at least about 0.1% of the active compound or more, although the percentage of the active ingredient(s) may, of course, be varied and may conveniently be between about 1 or 2% and about 60% or 70% or more of the weight or volume of the total formulation. Naturally, the amount of active compound(s) in each therapeutically useful composition may be prepared is such a way that a suitable dosage will be obtained in any given unit dose of the compound. Factors such as solubility, bioavailability, biological half-life, route of administration, product shelf life, as well as other pharmacological considerations will be contemplated by one skilled in the art of preparing such pharmaceutical formulations, and as such, a variety of dosages and treatment regimens may be desirable.

For oral administration the compositions of the present invention may alternatively be incorporated with one or more excipients in the form of a mouthwash, dentifrice, buccal tablet, oral spray, or sublingual orally-administered formulation. Alternatively, the active ingredient may be incorporated into an oral solution such as one containing sodium borate, glycerin and potassium bicarbonate, or dispersed in a dentifrice, or added in a therapeutically-effective amount to a composition that may include water, binders, abrasives, flavoring agents, foaming agents, and humectants. Alternatively the compositions may be fashioned into a tablet or solution form that may be placed under the tongue or otherwise dissolved in the mouth.

In certain circumstances it will be desirable to deliver the pharmaceutical compositions disclosed herein parenterally, intravenously, intramuscularly, or even

intraperitoneally. Such approaches are well known to the skilled artisan, some of which are further described, for example, in U. S. Patent 5,543,158; U. S. Patent 5,641,515 and U. S. Patent 5,399,363. In certain embodiments, solutions of the active compounds as free base or pharmacologically acceptable salts may be prepared in water suitably mixed with a surfactant, such as hydroxypropylcellulose. Dispersions may also be prepared in glycerol, liquid polyethylene glycols, and mixtures thereof and in oils. Under ordinary conditions of storage and use, these preparations generally will contain a preservative to prevent the growth of microorganisms.

Illustrative pharmaceutical forms suitable for injectable use include sterile aqueous solutions or dispersions and sterile powders for the extemporaneous preparation of sterile injectable solutions or dispersions (for example, see U. S. Patent 5,466,468). In all cases the form must be sterile and must be fluid to the extent that easy syringability exists. It must be stable under the conditions of manufacture and storage and must be preserved against the contaminating action of microorganisms, such as bacteria and fungi. The carrier can be a solvent or dispersion medium containing, for example, water, ethanol, polyol (*e.g.*, glycerol, propylene glycol, and liquid polyethylene glycol, and the like), suitable mixtures thereof, and/or vegetable oils. Proper fluidity may be maintained, for example, by the use of a coating, such as lecithin, by the maintenance of the required particle size in the case of dispersion and/or by the use of surfactants. The prevention of the action of microorganisms can be facilitated by various antibacterial and antifungal agents, for example, parabens, chlorobutanol, phenol, sorbic acid, thimerosal, and the like. In many cases, it will be preferable to include isotonic agents, for example, sugars or sodium chloride. Prolonged absorption of the injectable compositions can be brought about by the use in the compositions of agents delaying absorption, for example, aluminum monostearate and gelatin.

In one embodiment, for parenteral administration in an aqueous solution, the solution should be suitably buffered if necessary and the liquid diluent first rendered isotonic with sufficient saline or glucose. These particular aqueous solutions are especially suitable for intravenous, intramuscular, subcutaneous and intraperitoneal administration. In this connection, a sterile aqueous medium that can be employed will

be known to those of skill in the art in light of the present disclosure. For example, one dosage may be dissolved in 1 ml of isotonic NaCl solution and either added to 1000 ml of hypodermoclysis fluid or injected at the proposed site of infusion, (see for example, "Remington's Pharmaceutical Sciences" 15th Edition, pages 1035-1038 and 1570-1580). Some variation in dosage will necessarily occur depending on the condition of the subject being treated. Moreover, for human administration, preparations will of course preferably meet sterility, pyrogenicity, and the general safety and purity standards as required by FDA Office of Biologics standards.

In another embodiment of the invention, the compositions disclosed herein may be formulated in a neutral or salt form. Illustrative pharmaceutically-acceptable salts include the acid addition salts (formed with the free amino groups of the protein) and which are formed with inorganic acids such as, for example, hydrochloric or phosphoric acids, or such organic acids as acetic, oxalic, tartaric, mandelic, and the like. Salts formed with the free carboxyl groups can also be derived from inorganic bases such as, for example, sodium, potassium, ammonium, calcium, or ferric hydroxides, and such organic bases as isopropylamine, trimethylamine, histidine, procaine and the like. Upon formulation, solutions will be administered in a manner compatible with the dosage formulation and in such amount as is therapeutically effective.

The carriers can further comprise any and all solvents, dispersion media, vehicles, coatings, diluents, antibacterial and antifungal agents, isotonic and absorption delaying agents, buffers, carrier solutions, suspensions, colloids, and the like. The use of such media and agents for pharmaceutical active substances is well known in the art. Except insofar as any conventional media or agent is incompatible with the active ingredient, its use in the therapeutic compositions is contemplated. Supplementary active ingredients can also be incorporated into the compositions. The phrase "pharmaceutically-acceptable" refers to molecular entities and compositions that do not produce an allergic or similar untoward reaction when administered to a human.

In certain embodiments, the pharmaceutical compositions may be delivered by intranasal sprays, inhalation, and/or other aerosol delivery vehicles. Methods for delivering genes, nucleic acids, and peptide compositions directly to the

lungs *via* nasal aerosol sprays has been described, *e.g.,* in U. S. Patent 5,756,353 and U. S. Patent 5,804,212. Likewise, the delivery of drugs using intranasal microparticle resins (Takenaga *et al.*, J Controlled Release 1998 Mar 2;52(1-2):81-7) and lysophosphatidyl-glycerol compounds (U. S. Patent 5,725,871) are also well-known in the pharmaceutical arts. Likewise, illustrative transmucosal drug delivery in the form of a polytetrafluoroetheylene support matrix is described in U. S. Patent 5,780,045.

In certain embodiments, liposomes, nanocapsules, microparticles, lipid particles, vesicles, and the like, are used for the introduction of the compositions of the present invention into suitable host cells/organisms. In particular, the compositions of the present invention may be formulated for delivery either encapsulated in a lipid particle, a liposome, a vesicle, a nanosphere, or a nanoparticle or the like. Alternatively, compositions of the present invention can be bound, either covalently or non-covalently, to the surface of such carrier vehicles.

The formation and use of liposome and liposome-like preparations as potential drug carriers is generally known to those of skill in the art (see for example, Lasic, Trends Biotechnol 1998 Jul;16(7):307-21; Takakura, Nippon Rinsho 1998 Mar;56(3):691-5; Chandran *et al.*, Indian J Exp Biol. 1997 Aug;35(8):801-9; Margalit, Crit Rev Ther Drug Carrier Syst. 1995;12(2-3):233-61; U.S. Patent 5,567,434; U.S. Patent 5,552,157; U.S. Patent 5,565,213; U.S. Patent 5,738,868 and U.S. Patent 5,795,587, each specifically incorporated herein by reference in its entirety).

Liposomes have been used successfully with a number of cell types that are normally difficult to transfect by other procedures, including T cell suspensions, primary hepatocyte cultures and PC 12 cells (Renneisen *et al.,* J Biol Chem. 1990 Sep 25;265(27):16337-42; Muller *et al.*, DNA Cell Biol. 1990 Apr;9(3):221-9). In addition, liposomes are free of the DNA length constraints that are typical of viral-based delivery systems. Liposomes have been used effectively to introduce genes, various drugs, radiotherapeutic agents, enzymes, viruses, transcription factors, allosteric effectors and the like, into a variety of cultured cell lines and animals. Furthermore, he use of liposomes does not appear to be associated with autoimmune responses or unacceptable toxicity after systemic delivery.

In certain embodiments, liposomes are formed from phospholipids that are dispersed in an aqueous medium and spontaneously form multilamellar concentric bilayer vesicles (also termed multilamellar vesicles (MLVs).

Alternatively, in other embodiments, the invention provides for pharmaceutically-acceptable nanocapsule formulations of the compositions of the present invention. Nanocapsules can generally entrap compounds in a stable and reproducible way (see, for example, Quintanar-Guerrero *et al.*, Drug Dev Ind Pharm. 1998 Dec;24(12):1113-28). To avoid side effects due to intracellular polymeric overloading, such ultrafine particles (sized around 0.1 μm) may be designed using polymers able to be degraded *in vivo*. Such particles can be made as described, for example, by Couvreur *et al.,* Crit Rev Ther Drug Carrier Syst. 1988;5(1):1-20; zur Muhlen *et al.*, Eur J Pharm Biopharm. 1998 Mar;45(2):149-55; Zambaux *et al.* J Controlled Release. 1998 Jan 2;50(1-3):31-40; and U. S. Patent 5,145,684.

CANCER THERAPEUTIC METHODS

Immunologic approaches to cancer therapy are based on the recognition that cancer cells can often evade the body's defenses against aberrant or foreign cells and molecules, and that these defenses might be therapeutically stimulated to regain the lost ground, *e.g.* pgs. 623-648 in Klein, Immunology (Wiley-Interscience, New York, 1982). Numerous recent observations that various immune effectors can directly or indirectly inhibit growth of tumors has led to renewed interest in this approach to cancer therapy, *e.g.* Jager, et al., Oncology 2001;60(1):1-7; Renner, et al., Ann Hematol 2000 Dec;79(12):651-9.

Four-basic cell types whose function has been associated with antitumor cell immunity and the elimination of tumor cells from the body are: i) B-lymphocytes which secrete immunoglobulins into the blood plasma for identifying and labeling the nonself invader cells; ii) monocytes which secrete the complement proteins that are responsible for lysing and processing the immunoglobulin-coated target invader cells; iii) natural killer lymphocytes having two mechanisms for the destruction of tumor cells, antibody-dependent cellular cytotoxicity and natural killing; and iv) T-lymphocytes possessing antigen-specific receptors and having the capacity to recognize

a tumor cell carrying complementary marker molecules (Schreiber, H., 1989, in Fundamental Immunology (ed). W. E. Paul, pp. 923-955).

Cancer immunotherapy generally focuses on inducing humoral immune responses, cellular immune responses, or both. Moreover, it is well established that induction of $CD4^+$ T helper cells is necessary in order to secondarily induce either antibodies or cytotoxic $CD8^+$ T cells. Polypeptide antigens that are selective or ideally specific for cancer cells, particularly colon cancer cells, offer a powerful approach for inducing immune responses against colon cancer, and are an important aspect of the present invention.

Therefore, in further aspects of the present invention, the pharmaceutical compositions described herein may be used to stimulate an immune response against cancer, particularly for the immunotherapy of colon cancer. Within such methods, the pharmaceutical compositions described herein are administered to a patient, typically a warm-blooded animal, preferably a human. A patient may or may not be afflicted with cancer. Pharmaceutical compositions and vaccines may be administered either prior to or following surgical removal of primary tumors and/or treatment such as administration of radiotherapy or conventional chemotherapeutic drugs. As discussed above, administration of the pharmaceutical compositions may be by any suitable method, including administration by intravenous, intraperitoneal, intramuscular, subcutaneous, intranasal, intradermal, anal, vaginal, topical and oral routes.

Within certain embodiments, immunotherapy may be active immunotherapy, in which treatment relies on the *in vivo* stimulation of the endogenous host immune system to react against tumors with the administration of immune response-modifying agents (such as polypeptides and polynucleotides as provided herein).

Within other embodiments, immunotherapy may be passive immunotherapy, in which treatment involves the delivery of agents with established tumor-immune reactivity (such as effector cells or antibodies) that can directly or indirectly mediate antitumor effects and does not necessarily depend on an intact host immune system. Examples of effector cells include T cells as discussed above, T lymphocytes (such as $CD8^+$ cytotoxic T lymphocytes and $CD4^+$ T-helper tumor-

infiltrating lymphocytes), killer cells (such as Natural Killer cells and lymphokine-activated killer cells), B cells and antigen-presenting cells (such as dendritic cells and macrophages) expressing a polypeptide provided herein. T cell receptors and antibody receptors specific for the polypeptides recited herein may be cloned, expressed and transferred into other vectors or effector cells for adoptive immunotherapy. The polypeptides provided herein may also be used to generate antibodies or anti-idiotypic antibodies (as described above and in U.S. Patent No. 4,918,164) for passive immunotherapy.

Monoclonal antibodies may be labeled with any of a variety of labels for desired selective usages in detection, diagnostic assays or therapeutic applications (as described in U.S. Patent Nos. 6,090,365; 6,015,542; 5,843,398; 5,595,721; and 4,708,930, hereby incorporated by reference in their entirety as if each was incorporated individually). In each case, the binding of the labelled monoclonal antibody to the determinant site of the antigen will signal detection or delivery of a particular therapeutic agent to the antigenic determinant on the non-normal cell. A further object of this invention is to provide the specific monoclonal antibody suitably labelled for achieving such desired selective usages thereof.

Effector cells may generally be obtained in sufficient quantities for adoptive immunotherapy by growth *in vitro*, as described herein. Culture conditions for expanding single antigen-specific effector cells to several billion in number with retention of antigen recognition *in vivo* are well known in the art. Such *in vitro* culture conditions typically use intermittent stimulation with antigen, often in the presence of cytokines (such as IL-2) and non-dividing feeder cells. As noted above, immunoreactive polypeptides as provided herein may be used to rapidly expand antigen-specific T cell cultures in order to generate a sufficient number of cells for immunotherapy. In particular, antigen-presenting cells, such as dendritic, macrophage, monocyte, fibroblast and/or B cells, may be pulsed with immunoreactive polypeptides or transfected with one or more polynucleotides using standard techniques well known in the art. For example, antigen-presenting cells can be transfected with a polynucleotide having a promoter appropriate for increasing expression in a recombinant virus or other expression system. Cultured effector cells for use in therapy

must be able to grow and distribute widely, and to survive long term *in vivo*. Studies have shown that cultured effector cells can be induced to grow in vivo and to survive long term in substantial numbers by repeated stimulation with antigen supplemented with IL-2 (*see*, for example, Cheever et al., *Immunological Reviews 157*:177, 1997).

Alternatively, a vector expressing a polypeptide recited herein may be introduced into antigen presenting cells taken from a patient and clonally propagated *ex vivo* for transplant back into the same patient. Transfected cells may be reintroduced into the patient using any means known in the art, preferably in sterile form by intravenous, intracavitary, intraperitoneal or intratumor administration.

Routes and frequency of administration of the therapeutic compositions described herein, as well as dosage, will vary from individual to individual, and may be readily established using standard techniques. In general, the pharmaceutical compositions and vaccines may be administered by injection (*e.g.*, intracutaneous, intramuscular, intravenous or subcutaneous), intranasally (*e.g.*, by aspiration) or orally. Preferably, between 1 and 10 doses may be administered over a 52 week period. Preferably, 6 doses are administered, at intervals of 1 month, and booster vaccinations may be given periodically thereafter. Alternate protocols may be appropriate for individual patients. A suitable dose is an amount of a compound that, when administered as described above, is capable of promoting an anti-tumor immune response, and is at least 10-50% above the basal (*i.e.*, untreated) level. Such response can be monitored by measuring the anti-tumor antibodies in a patient or by vaccine-dependent generation of cytolytic effector cells capable of killing the patient's tumor cells *in vitro*. Such vaccines should also be capable of causing an immune response that leads to an improved clinical outcome (*e.g.*, more frequent remissions, complete or partial or longer disease-free survival) in vaccinated patients as compared to non-vaccinated patients. In general, for pharmaceutical compositions and vaccines comprising one or more polypeptides, the amount of each polypeptide present in a dose ranges from about 25 μg to 5 mg per kg of host. Suitable dose sizes will vary with the size of the patient, but will typically range from about 0.1 mL to about 5 mL.

In general, an appropriate dosage and treatment regimen provides the active compound(s) in an amount sufficient to provide therapeutic and/or prophylactic

benefit. Such a response can be monitored by establishing an improved clinical outcome (*e.g.*, more frequent remissions, complete or partial, or longer disease-free survival) in treated patients as compared to non-treated patients. Increases in preexisting immune responses to a tumor protein generally correlate with an improved clinical outcome. Such immune responses may generally be evaluated using standard proliferation, cytotoxicity or cytokine assays, which may be performed using samples obtained from a patient before and after treatment.

CANCER DETECTION AND DIAGNOSTIC COMPOSITIONS, METHODS AND KITS

In general, a cancer may be detected in a patient based on the presence of one or more colon tumor proteins and/or polynucleotides encoding such proteins in a biological sample (for example, blood, sera, sputum urine and/or tumor biopsies) obtained from the patient. In other words, such proteins may be used as markers to indicate the presence or absence of a cancer such as colon cancer. In addition, such proteins may be useful for the detection of other cancers. The binding agents provided herein generally permit detection of the level of antigen that binds to the agent in the biological sample.

Polynucleotide primers and probes may be used to detect the level of mRNA encoding a tumor protein, which is also indicative of the presence or absence of a cancer. In general, a tumor sequence should be present at a level that is at least two-fold, preferably three-fold, and more preferably five-fold or higher in tumor tissue than in normal tissue of the same type from which the tumor arose. Expression levels of a particular tumor sequence in tissue types different from that in which the tumor arose are irrelevant in certain diagnostic embodiments since the presence of tumor cells can be confirmed by observation of predetermined differential expression levels, e.g., 2-fold, 5-fold, etc, in tumor tissue to expression levels in normal tissue of the same type.

Other differential expression patterns can be utilized advantageously for diagnostic purposes. For example, in one aspect of the invention, overexpression of a tumor sequence in tumor tissue and normal tissue of the same type, but not in other normal tissue types, *e.g.* PBMCs, can be exploited diagnostically. In this case, the presence of metastatic tumor cells, for example in a sample taken from the circulation

or some other tissue site different from that in which the tumor arose, can be identified and/or confirmed by detecting expression of the tumor sequence in the sample, for example using RT-PCR analysis. In many instances, it will be desired to enrich for tumor cells in the sample of interest, e.g., PBMCs, using cell capture or other like techniques.

There are a variety of assay formats known to those of ordinary skill in the art for using a binding agent to detect polypeptide markers in a sample. *See, e.g.*, Harlow and Lane, *Antibodies: A Laboratory Manual,* Cold Spring Harbor Laboratory, 1988. In general, the presence or absence of a cancer in a patient may be determined by (a) contacting a biological sample obtained from a patient with a binding agent; (b) detecting in the sample a level of polypeptide that binds to the binding agent; and (c) comparing the level of polypeptide with a predetermined cut-off value.

In a preferred embodiment, the assay involves the use of binding agent immobilized on a solid support to bind to and remove the polypeptide from the remainder of the sample. The bound polypeptide may then be detected using a detection reagent that contains a reporter group and specifically binds to the binding agent/polypeptide complex. Such detection reagents may comprise, for example, a binding agent that specifically binds to the polypeptide or an antibody or other agent that specifically binds to the binding agent, such as an anti-immunoglobulin, protein G, protein A or a lectin. Alternatively, a competitive assay may be utilized, in which a polypeptide is labeled with a reporter group and allowed to bind to the immobilized binding agent after incubation of the binding agent with the sample. The extent to which components of the sample inhibit the binding of the labeled polypeptide to the binding agent is indicative of the reactivity of the sample with the immobilized binding agent. Suitable polypeptides for use within such assays include full length colon tumor proteins and polypeptide portions thereof to which the binding agent binds, as described above.

The solid support may be any material known to those of ordinary skill in the art to which the tumor protein may be attached. For example, the solid support may be a test well in a microtiter plate or a nitrocellulose or other suitable membrane. Alternatively, the support may be a bead or disc, such as glass, fiberglass, latex or a

plastic material such as polystyrene or polyvinylchloride. The support may also be a magnetic particle or a fiber optic sensor, such as those disclosed, for example, in U.S. Patent No. 5,359,681. The binding agent may be immobilized on the solid support using a variety of techniques known to those of skill in the art, which are amply described in the patent and scientific literature. In the context of the present invention, the term "immobilization" refers to both noncovalent association, such as adsorption, and covalent attachment (which may be a direct linkage between the agent and functional groups on the support or may be a linkage by way of a cross-linking agent). Immobilization by adsorption to a well in a microtiter plate or to a membrane is preferred. In such cases, adsorption may be achieved by contacting the binding agent, in a suitable buffer, with the solid support for a suitable amount of time. The contact time varies with temperature, but is typically between about 1 hour and about 1 day. In general, contacting a well of a plastic microtiter plate (such as polystyrene or polyvinylchloride) with an amount of binding agent ranging from about 10 ng to about 10 μg, and preferably about 100 ng to about 1 μg, is sufficient to immobilize an adequate amount of binding agent.

Covalent attachment of binding agent to a solid support may generally be achieved by first reacting the support with a bifunctional reagent that will react with both the support and a functional group, such as a hydroxyl or amino group, on the binding agent. For example, the binding agent may be covalently attached to supports having an appropriate polymer coating using benzoquinone or by condensation of an aldehyde group on the support with an amine and an active hydrogen on the binding partner (*see, e.g.*, Pierce Immunotechnology Catalog and Handbook, 1991, at A12-A13).

In certain embodiments, the assay is a two-antibody sandwich assay. This assay may be performed by first contacting an antibody that has been immobilized on a solid support, commonly the well of a microtiter plate, with the sample, such that polypeptides within the sample are allowed to bind to the immobilized antibody. Unbound sample is then removed from the immobilized polypeptide-antibody complexes and a detection reagent (preferably a second antibody capable of binding to a different site on the polypeptide) containing a reporter group is added. The amount of

detection reagent that remains bound to the solid support is then determined using a method appropriate for the specific reporter group.

More specifically, once the antibody is immobilized on the support as described above, the remaining protein binding sites on the support are typically blocked. Any suitable blocking agent known to those of ordinary skill in the art, such as bovine serum albumin or Tween 20™ (Sigma Chemical Co., St. Louis, MO). The immobilized antibody is then incubated with the sample, and polypeptide is allowed to bind to the antibody. The sample may be diluted with a suitable diluent, such as phosphate-buffered saline (PBS) prior to incubation. In general, an appropriate contact time (*i.e.*, incubation time) is a period of time that is sufficient to detect the presence of polypeptide within a sample obtained from an individual with colon cancer at least about 95% of that achieved at equilibrium between bound and unbound polypeptide. Those of ordinary skill in the art will recognize that the time necessary to achieve equilibrium may be readily determined by assaying the level of binding that occurs over a period of time. At room temperature, an incubation time of about 30 minutes is generally sufficient.

Unbound sample may then be removed by washing the solid support with an appropriate buffer, such as PBS containing 0.1% Tween 20™. The second antibody, which contains a reporter group, may then be added to the solid support. Preferred reporter groups include those groups recited above.

The detection reagent is then incubated with the immobilized antibody-polypeptide complex for an amount of time sufficient to detect the bound polypeptide. An appropriate amount of time may generally be determined by assaying the level of binding that occurs over a period of time. Unbound detection reagent is then removed and bound detection reagent is detected using the reporter group. The method employed for detecting the reporter group depends upon the nature of the reporter group. For radioactive groups, scintillation counting or autoradiographic methods are generally appropriate. Spectroscopic methods may be used to detect dyes, luminescent groups and fluorescent groups. Biotin may be detected using avidin, coupled to a different reporter group (commonly a radioactive or fluorescent group or an enzyme). Enzyme reporter groups may generally be detected by the addition of substrate

(generally for a specific period of time), followed by spectroscopic or other analysis of the reaction products.

To determine the presence or absence of a cancer, such as colon cancer, the signal detected from the reporter group that remains bound to the solid support is generally compared to a signal that corresponds to a predetermined cut-off value. In one preferred embodiment, the cut-off value for the detection of a cancer is the average mean signal obtained when the immobilized antibody is incubated with samples from patients without the cancer. In general, a sample generating a signal that is three standard deviations above the predetermined cut-off value is considered positive for the cancer. In an alternate preferred embodiment, the cut-off value is determined using a Receiver Operator Curve, according to the method of Sackett et al., *Clinical Epidemiology: A Basic Science for Clinical Medicine*, Little Brown and Co., 1985, p. 106-7. Briefly, in this embodiment, the cut-off value may be determined from a plot of pairs of true positive rates (*i.e.*, sensitivity) and false positive rates (100%-specificity) that correspond to each possible cut-off value for the diagnostic test result. The cut-off value on the plot that is the closest to the upper left-hand corner (*i.e.*, the value that encloses the largest area) is the most accurate cut-off value, and a sample generating a signal that is higher than the cut-off value determined by this method may be considered positive. Alternatively, the cut-off value may be shifted to the left along the plot, to minimize the false positive rate, or to the right, to minimize the false negative rate. In general, a sample generating a signal that is higher than the cut-off value determined by this method is considered positive for a cancer.

In a related embodiment, the assay is performed in a flow-through or strip test format, wherein the binding agent is immobilized on a membrane, such as nitrocellulose. In the flow-through test, polypeptides within the sample bind to the immobilized binding agent as the sample passes through the membrane. A second, labeled binding agent then binds to the binding agent-polypeptide complex as a solution containing the second binding agent flows through the membrane. The detection of bound second binding agent may then be performed as described above. In the strip test format, one end of the membrane to which binding agent is bound is immersed in a solution containing the sample. The sample migrates along the membrane through a

region containing second binding agent and to the area of immobilized binding agent. Concentration of second binding agent at the area of immobilized antibody indicates the presence of a cancer. Typically, the concentration of second binding agent at that site generates a pattern, such as a line, that can be read visually. The absence of such a pattern indicates a negative result. In general, the amount of binding agent immobilized on the membrane is selected to generate a visually discernible pattern when the biological sample contains a level of polypeptide that would be sufficient to generate a positive signal in the two-antibody sandwich assay, in the format discussed above. Preferred binding agents for use in such assays are antibodies and antigen-binding fragments thereof. Preferably, the amount of antibody immobilized on the membrane ranges from about 25 ng to about 1μg, and more preferably from about 50 ng to about 500 ng. Such tests can typically be performed with a very small amount of biological sample.

Of course, numerous other assay protocols exist that are suitable for use with the tumor proteins or binding agents of the present invention. The above descriptions are intended to be exemplary only. For example, it will be apparent to those of ordinary skill in the art that the above protocols may be readily modified to use tumor polypeptides to detect antibodies that bind to such polypeptides in a biological sample. The detection of such tumor protein specific antibodies may correlate with the presence of a cancer.

A cancer may also, or alternatively, be detected based on the presence of T cells that specifically react with a tumor protein in a biological sample. Within certain methods, a biological sample comprising $CD4^+$ and/or $CD8^+$ T cells isolated from a patient is incubated with a tumor polypeptide, a polynucleotide encoding such a polypeptide and/or an APC that expresses at least an immunogenic portion of such a polypeptide, and the presence or absence of specific activation of the T cells is detected. Suitable biological samples include, but are not limited to, isolated T cells. For example, T cells may be isolated from a patient by routine techniques (such as by Ficoll/Hypaque density gradient centrifugation of peripheral blood lymphocytes). T cells may be incubated *in vitro* for 2-9 days (typically 4 days) at 37°C with polypeptide (*e.g.*, 5 - 25 μg/ml). It may be desirable to incubate another aliquot of a T cell sample

in the absence of tumor polypeptide to serve as a control. For $CD4^+$ T cells, activation is preferably detected by evaluating proliferation of the T cells. For $CD8^+$ T cells, activation is preferably detected by evaluating cytolytic activity. A level of proliferation that is at least two fold greater and/or a level of cytolytic activity that is at least 20% greater than in disease-free patients indicates the presence of a cancer in the patient.

As noted above, a cancer may also, or alternatively, be detected based on the level of mRNA encoding a tumor protein in a biological sample. For example, at least two oligonucleotide primers may be employed in a polymerase chain reaction (PCR) based assay to amplify a portion of a tumor cDNA derived from a biological sample, wherein at least one of the oligonucleotide primers is specific for (*i.e.*, hybridizes to) a polynucleotide encoding the tumor protein. The amplified cDNA is then separated and detected using techniques well known in the art, such as gel electrophoresis.

Similarly, oligonucleotide probes that specifically hybridize to a polynucleotide encoding a tumor protein may be used in a hybridization assay to detect the presence of polynucleotide encoding the tumor protein in a biological sample.

To permit hybridization under assay conditions, oligonucleotide primers and probes should comprise an oligonucleotide sequence that has at least about 60%, preferably at least about 75% and more preferably at least about 90%, identity to a portion of a polynucleotide encoding a tumor protein of the invention that is at least 10 nucleotides, and preferably at least 20 nucleotides, in length. Preferably, oligonucleotide primers and/or probes hybridize to a polynucleotide encoding a polypeptide described herein under moderately stringent conditions, as defined above. Oligonucleotide primers and/or probes which may be usefully employed in the diagnostic methods described herein preferably are at least 10-40 nucleotides in length. In a preferred embodiment, the oligonucleotide primers comprise at least 10 contiguous nucleotides, more preferably at least 15 contiguous nucleotides, of a DNA molecule having a sequence as disclosed herein. Techniques for both PCR based assays and hybridization assays are well known in the art (*see*, for example, Mullis et al., *Cold*

Spring Harbor Symp. Quant. Biol., *51*:263, 1987; Erlich ed., *PCR Technology*, Stockton Press, NY, 1989).

One preferred assay employs RT-PCR, in which PCR is applied in conjunction with reverse transcription. Typically, RNA is extracted from a biological sample, such as biopsy tissue, and is reverse transcribed to produce cDNA molecules. PCR amplification using at least one specific primer generates a cDNA molecule, which may be separated and visualized using, for example, gel electrophoresis. Amplification may be performed on biological samples taken from a test patient and from an individual who is not afflicted with a cancer. The amplification reaction may be performed on several dilutions of cDNA spanning two orders of magnitude. A two-fold or greater increase in expression in several dilutions of the test patient sample as compared to the same dilutions of the non-cancerous sample is typically considered positive.

In another aspect of the present invention, cell capture technologies may be used in conjunction, with, for example, real-time PCR to provide a more sensitive tool for detection of metastatic cells expressing colon tumor antigens. Detection of colon cancer cells in biological samples, e.g., bone marrow samples, peripheral blood, and small needle aspiration samples is desirable for diagnosis and prognosis in colon cancer patients.

Immunomagnetic beads coated with specific monoclonal antibodies to surface cell markers, or tetrameric antibody complexes, may be used to first enrich or positively select cancer cells in a sample. Various commercially available kits may be used, including Dynabeads® Epithelial Enrich (Dynal Biotech, Oslo, Norway), StemSep™ (StemCell Technologies, Inc., Vancouver, BC), and RosetteSep (StemCell Technologies). A skilled artisan will recognize that other methodologies and kits may also be used to enrich or positively select desired cell populations. Dynabeads® Epithelial Enrich contains magnetic beads coated with mAbs specific for two glycoprotein membrane antigens expressed on normal and neoplastic epithelial tissues. The coated beads may be added to a sample and the sample then applied to a magnet, thereby capturing the cells bound to the beads. The unwanted cells are washed away and the magnetically isolated cells eluted from the beads and used in further analyses.

RosetteSep can be used to enrich cells directly from a blood sample and consists of a cocktail of tetrameric antibodies that targets a variety of unwanted cells and crosslinks them to glycophorin A on red blood cells (RBC) present in the sample, forming rosettes. When centrifuged over Ficoll, targeted cells pellet along with the free RBC. The combination of antibodies in the depletion cocktail determines which cells will be removed and consequently which cells will be recovered. Antibodies that are available include, but are not limited to: CD2, CD3, CD4, CD5, CD8, CD10, CD11b, CD14, CD15, CD16, CD19, CD20, CD24, CD25, CD29, CD33, CD34, CD36, CD38, CD41, CD45, CD45RA, CD45RO, CD56, CD66B, CD66e, HLA-DR, IgE, and TCR$\alpha\beta$.

Additionally, it is contemplated in the present invention that mAbs specific for colon tumor antigens can be generated and used in a similar manner. For example, mAbs that bind to tumor-specific cell surface antigens may be conjugated to magnetic beads, or formulated in a tetrameric antibody complex, and used to enrich or positively select metastatic colon tumor cells from a sample. Once a sample is enriched or positively selected, cells may be lysed and RNA isolated. RNA may then be subjected to RT-PCR analysis using colon tumor-specific primers in a real-time PCR assay as described herein. One skilled in the art will recognize that enriched or selected populations of cells may be analyzed by other methods (*e.g. in situ* hybridization or flow cytometry).

In another embodiment, the compositions described herein may be used as markers for the progression of cancer. In this embodiment, assays as described above for the diagnosis of a cancer may be performed over time, and the change in the level of reactive polypeptide(s) or polynucleotide(s) evaluated. For example, the assays may be performed every 24-72 hours for a period of 6 months to 1 year, and thereafter performed as needed. In general, a cancer is progressing in those patients in whom the level of polypeptide or polynucleotide detected increases over time. In contrast, the cancer is not progressing when the level of reactive polypeptide or polynucleotide either remains constant or decreases with time.

Certain *in vivo* diagnostic assays may be performed directly on a tumor. One such assay involves contacting tumor cells with a binding agent. The bound

binding agent may then be detected directly or indirectly via a reporter group. Such binding agents may also be used in histological applications. Alternatively, polynucleotide probes may be used within such applications.

As noted above, to improve sensitivity, multiple tumor protein markers may be assayed within a given sample. It will be apparent that binding agents specific for different proteins provided herein may be combined within a single assay. Further, multiple primers or probes may be used concurrently. The selection of tumor protein markers may be based on routine experiments to determine combinations that results in optimal sensitivity. In addition, or alternatively, assays for tumor proteins provided herein may be combined with assays for other known tumor antigens.

The present invention further provides kits for use within any of the above diagnostic methods. Such kits typically comprise two or more components necessary for performing a diagnostic assay. Components may be compounds, reagents, containers and/or equipment. For example, one container within a kit may contain a monoclonal antibody or fragment thereof that specifically binds to a tumor protein. Such antibodies or fragments may be provided attached to a support material, as described above. One or more additional containers may enclose elements, such as reagents or buffers, to be used in the assay. Such kits may also, or alternatively, contain a detection reagent as described above that contains a reporter group suitable for direct or indirect detection of antibody binding.

Alternatively, a kit may be designed to detect the level of mRNA encoding a tumor protein in a biological sample. Such kits generally comprise at least one oligonucleotide probe or primer, as described above, that hybridizes to a polynucleotide encoding a tumor protein. Such an oligonucleotide may be used, for example, within a PCR or hybridization assay. Additional components that may be present within such kits include a second oligonucleotide and/or a diagnostic reagent or container to facilitate the detection of a polynucleotide encoding a tumor protein.

The following Examples are offered by way of illustration and not by way of limitation.

EXAMPLES

EXAMPLE 1

GENERATION OF COLON ADENOCARCINOMA-SPECIFIC SUBTRACTED cDNA LIBRARIES

Colon tumor subtracted cDNA libraries were constructed. Briefly, a pool of tester mRNA was collected from three colon adenocarcinoma samples showing moderate histological differentiation and no evidence of metastasis. Eight normal tissues, including brain, pancreas, bone marrow, liver, heart, lung, stomach and small intestine were represented in the driver mRNA pool. cDNA synthesis, hybridization and PCR amplification were performed according to the methods of Clontech (Palo Alto, CA), with minor modifications. In a first subtraction, the restriction enzymes PvuII, DraI, MscI and StuI were used to digest cDNAs. The tester to driver ratio was 1:40. In a second subtraction, DraI, MscI and StuI were used for cDNA digestion. A tester to driver ratio of 1:76 was employed. Following the PCR amplification steps, the cDNAs were cloned into the pCR2.1 plasmid vector. The libraries resulting from the first and second subtractions, named CPS1 and CPS2, respectively, were used to obtain clones for microarray analysis and sequencing. Inserts were PCR amplified and purified. Each clone was sequenced from one direction with either M13 Forward primer or M13 Reverse primer.

In another subtraction, a cDNA library was constructed in the PCR2.1 vector (Invitrogen, Carlsbad, CA) by subtracting a pool of three colon tumors with a pool of normal colon, spleen, brain, liver, kidney, lung, stomach and small intestine using PCR subtraction methodologies (Clontech, Palo Alto, CA). The subtraction was performed using a PCR-based protocol, which was modified to generate larger fragments. Within this protocol, tester and driver double stranded cDNA were separately digested with five restriction enzymes that recognize six-nucleotide restriction sites (MluI, MscI, PvuII, SalI and StuI). This digestion resulted in an average cDNA size of 600 bp, rather than the average size of 300 bp that results from digestion with RsaI according to the Clontech protocol. This modification did not affect the subtraction efficiency. Two tester populations were then created with different adapters, and the driver library remained without adapters.

The tester and driver libraries were then hybridized using excess driver cDNA. In the first hybridization step, driver was separately hybridized with each of the two tester cDNA populations. This resulted in populations of (a) unhybridized tester cDNAs, (b) tester cDNAs hybridized to other tester cDNAs, (c) tester cDNAs hybridized to driver cDNAs, and (d) unhybridized driver cDNAs. The two separate hybridization reactions were then combined, and rehybridized in the presence of additional denatured driver cDNA. Following this second hybridization, in addition to populations (a) through (d), a fifth population (e) was generated in which tester cDNA with one adapter hybridized to tester cDNA with the second adapter. Accordingly, the second hybridization step resulted in enrichment of differentially expressed sequences which could be used as templates for PCR amplification with adaptor-specific primers.

The ends were then filled in, and PCR amplification was performed using adaptor-specific primers. Only population (e), which contained tester cDNA that did not hybridize to driver cDNA, was amplified exponentially. A second PCR amplification step was then performed, to reduce background and further enrich differentially expressed sequences. This PCR-based subtraction technique normalizes differentially expressed cDNAs so that rare transcripts that are over-expressed in colon tumor tissue may be recoverable. Such transcripts would be difficult to recover by traditional subtraction methods.

The determined cDNA sequences for 333 clones from the colon tumor subtracted libraries are provided in SEQ ID NO: 1-333.

EXAMPLE 2

ANALYSIS OF SUBTRACTED CDNA SEQUENCES BY MICROARRAY ANALYSIS

In additional studies, subtracted cDNA sequences were analyzed by microarray analysis to evaluate their expression in tumor and normal tissues. Using this approach, cDNA sequences are PCR amplified and their mRNA expression profiles in tumor and normal tissues are examined using cDNA microarray technology essentially as described (Shena *et al.*, 1995). In brief, the clones are arrayed onto glass slides as multiple replicas, with each location corresponding to a unique cDNA clone (as many as 5500 clones can be arrayed on a single slide, or chip). Each chip is

hybridized with a pair of cDNA probes that are fluorescence-labeled with Cy3 and Cy5, respectively. Typically, 1μg of polyA$^+$ RNA is used to generate each cDNA probe. After hybridization, the chips are scanned and the fluorescence intensity recorded for both Cy3 and Cy5 channels. There are multiple built-in quality control steps. First, the probe quality is monitored using a panel of ubiquitously expressed genes. Secondly, the control plate also can include yeast DNA fragments of which complementary RNA may be spiked into the probe synthesis for measuring the quality of the probe and the sensitivity of the analysis. Currently, the technology offers a sensitivity of 1 in 100,000 copies of mRNA. Finally, the reproducibility of this technology can be ensured by including duplicated control cDNA elements at different locations.

Analysis of colon tumor subtracted clones by microarray analysis on Colon Chip 3 identified the sequences set forth in SEQ ID NO:335-377 as being at least two-fold overexpressed in colon tumors versus normal tissues.

EXAMPLE 3

IDENTIFICATION OF NORMAL COLON cDNAs

Clones were derived from a characterization of a primary normal colon library. Two normal colon tissue samples were represented in the mRNA pool. These clones were sequenced and data base searches performed. SEQ ID NO:334 disclosed herein showed no homology to known sequences.

EXAMPLE 4

ANALYSIS OF C592S cDNA EXPRESSION USING REAL-TIME PCR

The colon tumor antigen, C592S (SEQ ID NO:348), was isolated from the subtraction library described in Example 1 and was found by microarray analysis to be overexpressed in colon tumors as compared to normal colon tissue. This sequence shows no significant similarity to known sequences in Genbank. The expression pattern of this gene was further analyzed by real-time PCR, as described below, and was found to be overexpressed in colon tumor while it was expressed at lower levels in normal colon. No expression was observed in a panel of other normal tissues. This

data indicates that C592S may be valuable as a tumor immunotherapeutic or diagnostic tool.

The first-strand cDNA to be used in the quantitative real-time PCR was synthesized from 20μg of total RNA that had been treated with DNase I (Amplification Grade, Gibco BRL Life Technology, Gaithersburg, MD), using Superscript Reverse Transcriptase (RT) (Gibco BRL Life Technology, Gaithersburg, MD). Real-time PCR was performed with a GeneAmp™ 5700 sequence detection system (PE Biosystems, Foster City, CA). The 5700 system uses SYBR™ green, a fluorescent dye that only intercalates into double stranded DNA, and a set of gene-specific forward and reverse primers. The increase in fluorescence is monitored during the whole amplification process. The optimal concentration of primers was determined using a checkerboard approach and a pool of cDNAs from breast tumors was used in this process. The PCR reaction was performed in 25μl volumes that include 2.5μl of SYBR green buffer, 2μl of cDNA template and 2.5μl each of the forward and reverse primers for the gene of interest. The cDNAs used for RT reactions were diluted 1:10 for each gene of interest and 1:100 for the β-actin control. In order to quantitate the amount of specific cDNA (and hence initial mRNA) in the sample, a standard curve is generated for each run using the plasmid DNA containing the gene of interest. Standard curves were generated using the Ct values determined in the real-time PCR which were related to the initial cDNA concentration used in the assay. Standard dilution ranging from 20-2x10^6 copies of the gene of interest was used for this purpose. In addition, a standard curve was generated for β-actin ranging from 200fg-2000fg. This enabled standardization of the initial RNA content of a tissue sample to the amount of β-actin for comparison purposes. The mean copy number for each group of tissues tested was normalized to a constant amount of β-actin, allowing the evaluation of the over-expression levels seen with each of the genes.

EXAMPLE 5

PEPTIDE PRIMING OF T-HELPER LINES

Generation of $CD4^+$ T helper lines and identification of peptide epitopes derived from tumor-specific antigens that are capable of being recognized by $CD4^+$ T cells in the context of HLA class II molecules, is carried out as follows:

Fifteen-mer peptides overlapping by 10 amino acids, derived from a tumor-specific antigen, are generated using standard procedures. Dendritic cells (DC) are derived from PBMC of a normal donor using GM-CSF and IL-4 by standard protocols. $CD4^+$ T cells are generated from the same donor as the DC using MACS beads (Miltenyi Biotec, Auburn, CA) and negative selection. DC are pulsed overnight with pools of the 15-mer peptides, with each peptide at a final concentration of 0.25 μg/ml. Pulsed DC are washed and plated at 1 x 10^4 cells/well of 96-well V-bottom plates and purified $CD4^+$ T cells are added at 1 x 10^5/well. Cultures are supplemented with 60 ng/ml IL-6 and 10 ng/ml IL-12 and incubated at 37°C. Cultures are restimulated as above on a weekly basis using DC generated and pulsed as above as antigen presenting cells, supplemented with 5 ng/ml IL-7 and 10 U/ml IL-2. Following 4 *in vitro* stimulation cycles, resulting $CD4^+$ T cell lines (each line corresponding to one well) are tested for specific proliferation and cytokine production in response to the stimulating pools of peptide with an irrelevant pool of peptides used as a control.

EXAMPLE 6

GENERATION OF TUMOR-SPECIFIC CTL LINES USING IN VITRO WHOLE-GENE PRIMING

Using *in vitro* whole-gene priming with tumor antigen-vaccinia infected DC (see, for example, Yee et al, *The Journal of Immunology*, 157(9):4079-86, 1996), human CTL lines are derived that specifically recognize autologous fibroblasts transduced with a specific tumor antigen, as determined by interferon-γ ELISPOT analysis. Specifically, dendritic cells (DC) are differentiated from monocyte cultures derived from PBMC of normal human donors by growing for five days in RPMI medium containing 10% human serum, 50 ng/ml human GM-CSF and 30 ng/ml human IL-4. Following culture, DC are infected overnight with tumor antigen-recombinant vaccinia virus at a multiplicity of infection (M.O.I) of five, and matured overnight by

the addition of 3 μg/ml CD40 ligand. Virus is then inactivated by UV irradiation. CD8+ T cells are isolated using a magnetic bead system, and priming cultures are initiated using standard culture techniques. Cultures are restimulated every 7-10 days using autologous primary fibroblasts retrovirally transduced with previously identified tumor antigens. Following four stimulation cycles, CD8+ T cell lines are identified that specifically produce interferon-γ when stimulated with tumor antigen-transduced autologous fibroblasts. Using a panel of HLA-mismatched B-LCL lines transduced with a vector expressing a tumor antigen, and measuring interferon-γ production by the CTL lines in an ELISPOT assay, the HLA restriction of the CTL lines is determined.

EXAMPLE 7

GENERATION AND CHARACTERIZATION OF ANTI-TUMOR ANTIGEN MONOCLONAL ANTIBODIES

Mouse monoclonal antibodies are raised against *E. coli* derived tumor antigen proteins as follows: Mice are immunized with Complete Freund's Adjuvant (CFA) containing 50 μg recombinant tumor protein, followed by a subsequent intraperitoneal boost with Incomplete Freund's Adjuvant (IFA) containing 10μg recombinant protein. Three days prior to removal of the spleens, the mice are immunized intravenously with approximately 50μg of soluble recombinant protein. The spleen of a mouse with a positive titer to the tumor antigen is removed, and a single-cell suspension made and used for fusion to SP2/O myeloma cells to generate B cell hybridomas. The supernatants from the hybrid clones are tested by ELISA for specificity to recombinant tumor protein, and epitope mapped using peptides that spanned the entire tumor protein sequence. The mAbs are also tested by flow cytometry for their ability to detect tumor protein on the surface of cells stably transfected with the cDNA encoding the tumor protein.

CLAIMS

What is claimed:

1. An isolated polynucleotide comprising a sequence selected from the group consisting of:

(a) sequences provided in SEQ ID NO:1-377;

(b) complements of the sequences provided in SEQ ID NO:1-377;

(c) sequences consisting of at least 20 contiguous residues of a sequence provided in SEQ ID NO:1-377;

(d) sequences that hybridize to a sequence provided in SEQ ID NO:1-377, under highly stringent conditions;

(e) sequences having at least 75% identity to a sequence of SEQ ID NO:1-377;

(f) sequences having at least 90% identity to a sequence of SEQ ID NO:1-377; and

(g) degenerate variants of a sequence provided in SEQ ID NO:1-377.

2. An isolated polypeptide comprising an amino acid sequence selected from the group consisting of:

(a) sequences encoded by a polynucleotide of claim 1; and

(b) sequences having at least 70% identity to a sequence encoded by a polynucleotide of claim 1; and

(c) sequences having at least 90% identity to a sequence encoded by a polynucleotide of claim 1.

3. An expression vector comprising a polynucleotide of claim 1 operably linked to an expression control sequence.

4. A host cell transformed or transfected with an expression vector according to claim 3.

5. An isolated antibody, or antigen-binding fragment thereof, that specifically binds to a polypeptide of claim 2.

6. A method for detecting the presence of a cancer in a patient, comprising the steps of:

(a) obtaining a biological sample from the patient;

(b) contacting the biological sample with a binding agent that binds to a polypeptide of claim 2;

(c) detecting in the sample an amount of polypeptide that binds to the binding agent; and

(d) comparing the amount of polypeptide to a predetermined cut-off value and therefrom determining the presence of a cancer in the patient.

7. A fusion protein comprising at least one polypeptide according to claim 2.

8. An oligonucleotide that hybridizes to a sequence recited in SEQ ID NO:1-377 under highly stringent conditions.

9. A method for stimulating and/or expanding T cells specific for a tumor protein, comprising contacting T cells with at least one component selected from the group consisting of:

(a) polypeptides according to claim 2;

(b) polynucleotides according to claim 1; and

(c) antigen-presenting cells that express a polynucleotide according to claim 1,

under conditions and for a time sufficient to permit the stimulation and/or expansion of T cells.

10. An isolated T cell population, comprising T cells prepared according to the method of claim 9.

11. A composition comprising a first component selected from the group consisting of physiologically acceptable carriers and immunostimulants, and a second component selected from the group consisting of:

(a) polypeptides according to claim 2;

(b) polynucleotides according to claim 1;

(c) antibodies according to claim 5;

(d) fusion proteins according to claim 7;

(e) T cell populations according to claim 10; and

(f) antigen presenting cells that express a polypeptide according to claim 2.

12. A method for stimulating an immune response in a patient, comprising administering to the patient a composition of claim 11.

13. A method for the treatment of a colon cancer in a patient, comprising administering to the patient a composition of claim 11.

14. A method for determining the presence of a cancer in a patient, comprising the steps of:

(a) obtaining a biological sample from the patient;

(b) contacting the biological sample with an oligonucleotide according to claim 8;

(c) detecting in the sample an amount of a polynucleotide that hybridizes to the oligonucleotide; and

(d) compare the amount of polynucleotide that hybridizes to the oligonucleotide to a predetermined cut-off value, and therefrom determining the presence of the cancer in the patient.

15. A diagnostic kit comprising at least one oligonucleotide according to claim 8.

16. A diagnostic kit comprising at least one antibody according to claim 5 and a detection reagent, wherein the detection reagent comprises a reporter group.

17. A method for the treatment of colon cancer in a patient, comprising the steps of:

(a) incubating CD4+ and/or CD8+ T cells isolated from a patient with at least one component selected from the group consisting of: (i) polypeptides according to claim 2; (ii) polynucleotides according to claim 1; and (iii) antigen presenting cells that express a polypeptide of claim 2, such that T cell proliferate;

(b) administering to the patient an effective amount of the proliferated T cells, and thereby inhibiting the development of a cancer in the patient.

SEQUENCE LISTING

<110> Corixa Coporation
Meagher, Madeleine Joy
King, Gordon E.
Xu, Jiangchun
Secrist, Heather

<120> COMPOSITIONS AND METHODS FOR THE
THERAPY AND DIAGNOSIS OF COLON CANCER

<130> 210121.532PC

<140> PCT
<141> 2001-06-07

<160> 377

<170> Corixa Invention Disclosure Database

<210> 1
<211> 291
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(291)
<223> n = A,T,C or G

<400> 1
cgcggggaca aactacttca tcaaggtgca cgtcggcgac naggacttcg tacacctgcg 60
agtgttccaa tctctccctc atgaaaacaa gcccttgacc ttatctaact accagaccaa 120
caaagccaag catgatgagc tgacctattt ctgatcctga ctttggacaa ggcccttcag 180
ccagaanact gacaaagtca tcctccgtct accanancgt gcacttgtga tcctaaaata 240
agcttcatct ncgggctgtg ccccttgggg tggaaggggc aggattctgc a 291

<210> 2
<211> 549
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(549)
<223> n = A,T,C or G

<400> 2
gcttagtttt acattcacaa agtatttgta tttgataagt cttaaaaatg atttaagttg 60
aagcctaaca gaagtacttg ttcattcctc agttacctgc aaataatgga aataatggtg 120
cagtcagttt tatttgattg tgtaatgctt tatactaggg gtctcaaacc cctgggtggc 180
attagattct tacaggagcg tgaaccctat tgtgaactgc acatgcaaga gatctaggtt 240
gtgcgctcat tatgagaatc taatacccga tgatctgagg tggaacagtc tcatnctgaa 300

aaccatctac cccaacacct gtgtctgtgg aaattgtctc cacgaaaccg gtccctgttg 360
aaaaaaggtt aggggaccac tgctttatac agtgtttgaa gtaggcacac ttacagttac 420
tatttctggt tctttatgac tgttgagggg ggcagaattc tatgtttgat tcacagtaaa 480
acttttcatg aatcattgag gggttaactg ttatcacttg ccaatggaaa ntggatattg 540
aaacatttt 549

<210> 3
<211> 298
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(298)
<223> n = A,T,C or G

<400> 3
gtggtgcgcg cccgcnagca gatcgtagct ggctgctggc acanaaaggc atctattttt 60
caatggtcag tgtccaggct ggaacaaagc gccagtgaac tctgactgta taagatgtta 120
aatacttttt aatatttgtt tagatatgac atttattcaa agttaaaagc aaacactnnc 180
agaattatga agaggtatct gtttaacatt tnctcagtca agttcagagt cttcagagac 240
ttcgtaatta aaggaacaga gtgagagaca tcatcaagtg gagagaaatc ataggttt 298

<210> 4
<211> 218
<212> DNA
<213> Homo sapiens

<400> 4
ccgggcaggt aaaaaacctt tgaaaaatat acagcttgat attatttaca taaaatatga 60
atccaggttc caatatcaaa caaacattgc tatgtcagaa acacagtgga aggcaggaac 120
gtaactcact gccttttaga tgcaaagact aatagacacg ttctcctatc tcgactatct 180
ttgttacctg ttatcctcaa acataaatta ttaggcac 218

<210> 5
<211> 218
<212> DNA
<213> Homo sapiens

<400> 5
cctgacattc ctgccttctt ataataagaa aaataaaaca aaatagtgtt gaagtgttgg 60
ggcggcgaaa atttttgggg ggtggaatgg agagagaatg ggcgatgttt ctcagggctg 120
cttcaagcgg gattaggggc ggcgtgggaa cctagagtgg gagagattaa gctgaaggga 180
ggtcttgtgg taaggggtga tattgtgggg atgttaaa 218

<210> 6
<211> 495
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(495)
<223> n = A,T,C or G

<400> 6
aaaaagaaaa aaaaagccaa atacattttc tgacattgta agattgcctt actgtctgtc 60
attccttatt gctggcccct ttctcaggcc ggaggccaag tggtggagaa ggaaaggaaa 120

tgatcgaacg ggcatgttgt caagtgggca tgccactggg aaataccacc agtttaccct 180
gaaacattgt cctcagagga gtaggaaagt ggattttgaa tctctatttt gctcaaaagt 240
tcagttcctg agatactgat gactgagagt gctgctggga aattttcagg attgtgtggt 300
cttttggggt tttttgtttt tttttttaa gacaaagttg accgctgttc actgtccacg 360
tgatcagttg taagattaca atgctgcatg ctagttggtt acataagata caattccagt 420
gatggaaggc ggttataatg gatggtggtg tgtacaanat ggcactgcca tctttgagca 480
gagcccagct ctgca 495

<210> 7
<211> 363
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(363)
<223> n = A,T,C or G

<400> 7
ccttcggaca gtagataagt atccaaagca gtgtgggcaa aaagagtaag caacatcccc 60
tgggattttt cccatgtggg tatcactgag caaagctgct gtaggaagat atgtagatta 120
atgggaaatg cagggttatt atggaggaaa taactcacag gatgatttca tagcctggct 180
aaattactac ttgattaact cctgataaaa ctcagagcgc acaaagcgag gcagagaatc 240
cttttccatc agggcatgga ttcttttctg ggccatgtca aagctgctca nggaaaggtt 300
ccaccaggtt tcttcattgt gatgtcctta gtgaagtggt caatattcac ctctttagga 360
gcc 363

<210> 8
<211> 403
<212> DNA
<213> Homo sapiens

<400> 8
ctgtccaatg gcaacatgac cctcactcta ctcagcgtca aaaggaacga tgcaggatcc 60
tatgaatgtg aaatacagaa cccagcgagt gccaaccgca gtgacccagt caccctgaat 120
gtcctctatg gcccagatgg ccccaccatt tccccctcaa aggccaatta ccgtccaggg 180
gaaaatctga acctctcctg ccacgcagcc tctaacccac ctgcacagta ctcttggttt 240
atcaatggga cgttccagca atccacacaa gagctcttta tccccaacat cactgtgaat 300
aatagcggat cctatatgtg ccaagcccat aactcagcca ctggcctcaa taggaccaca 360
gtcacgatga ccacagtctc tggaagtgct cctgtcctct cag 403

<210> 9
<211> 331
<212> DNA
<213> Homo sapiens

<400> 9
tttttaactc ctctcgctct gatgggacat ttgttaccct tttttcatag tgaaattgtg 60
tttcaggctt agtctgacct ttctggtttc ttcattttct tccattactt aggaaagagt 120
ggaaactcca ctaaaatttc tctgtgttgt tacagtctta gaggttgcag tactatattg 180
taagctttgg tgtttgttta attagcaata gggatggtag gattcaaatg tgtgtcattt 240
agaagtggaa gctattagca ccaatgacat aaatacatac aagacacaca actaaaatgt 300
catgttatta acagttatta ggttgtcatt t 331

<210> 10
<211> 253
<212> DNA
<213> Homo sapiens

<400> 10
ccaggcccca ggtctcctat ttgggagaac cactgccctc tgcctgcctc tccaactact 60
gctgggctgc ggcccaggcg ccttcaacga ccattttagg gttctgatga aagcaccttc 120
ggcttctaag gtgcaggctg ggaaacaagg tgggggccca catagcctgg tgtctcagca 180
tggagcttag tgccaagtcc tgtgccagag acacctgatg tgtaaagagg gaagagggca 240
cacttgggag tgg 253

<210> 11
<211> 298
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(298)
<223> n = A,T,C or G

<400> 11
cctacagact tanctcttct tggacacacc cacggngcgg ccacggcngc cagtggtctt 60
ggtgtgctgg cctcngacac naaggcccca gaagtgacgc agccctctat gggcccgaat 120
cttcttcagt cgctccaggt cttcacggag cttgttgtcc agaccattgg ctaggacctg 180
gctgtatttt ccatccttta catccttctg tctgttcaag aaccagtctg ggatcttgta 240
ctggcgtgga ttctgcataa tggtgatcac acgttccacc tnatnctcag tgagttct 298

<210> 12
<211> 344
<212> DNA
<213> Homo sapiens

<400> 12
ctgtagtccc agttactcgg gaggctgagg caggagaatc gcttgaaccc gggaggtgga 60
gattgcagtg agcccagatc gcaccactgc actccagtct ggcaacagag caagactcca 120
tctcaaaaag aaaagaaaag aagactctga cctgtactct tgaatacaag tttctgatac 180
cactgcactg tctgagaatt tccaaaactt taatgaacta actgacagct tcatgaaact 240
gtccaccaag atcaagcaga gaaaataatt aatttcatgg gactaaagga actaatgagg 300
ataatatttt cataattttt tatttgaaat tttgctgatt cttt 344

<210> 13
<211> 230
<212> DNA
<213> Homo sapiens

<400> 13
ccttatttct cttgtccttt cgtacaggga ggaatttgaa gtagatagaa accgacctgg 60
attactccgg tctgaactca gatcacgtag gactttaatc gttgaacaaa cgaaccttta 120
atagcggctg caccatcggg atgtcctgat ccaacatcga ggtcgtaaac cctattgttg 180
atatggactc tagaatagga ttgcgctgtt atccctaggg taacttgttc 230

<210> 14
<211> 216
<212> DNA
<213> Homo sapiens

<400> 14
cctgacattc ctgccttctt ataataagaa aaataaaaca aaatagtgtt gaagtgttgg 60
ggcggcgaaa atttttgggg ggtggaatgg agagagaatg ggcgatgttt ctcagggctg 120
cttcaagcgg gattaggggc ggcgtgggaa cctagagtgg gagagattaa gctgaaggga 180

ggtcttgtgg taaggggtga tattgtgggg atgtta 216

<210> 15
<211> 159
<212> DNA
<213> Homo sapiens

<400> 15
ctggtggtga ttgcacacga cgtggatccc atcgagctgg ttgtcttctt gcctgccctg 60
tgtcgtaaaa tgggggtccc ttactgcatt atcaagggaa aggcaagact gggacgtcta 120
gtccacagga agacctgcac cactgtcgcc ttcacacag 159

<210> 16
<211> 462
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(462)
<223> n = A,T,C or G

<400> 16
ccatcacaag ccacgagggc tggatgatac cttaacatga gacagccaaa tgcttaggca 60
gataaaatgg ggtccctgga gaatctccaa gcgtcccaag aatgtttaca ttagatgctt 120
ttgggtcggt gagggaacct gcccagggct tgtctgggca tcccacagtg aactggagcc 180
tgacgtacgc actgggggaa gtgggtgggg ccacggggaa ttcttccacg gggaagagaa 240
gcctgctctc ttccgctcct gtagtgactg taccagccag accaggaaga ccctggggtc 300
cangggggac cacgttctcc actgagaccg ttagctcctg gtttcccact ttcacccttg 360
acaccctgag ggccagggct tcccctagga cctggcatgc ctggtggtcc tgcaagaccc 420
cgtgctccag tgatcccagc aatcccaagt gggcctggag ct 462

<210> 17
<211> 103
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(103)
<223> n = A,T,C or G

<400> 17
cgccccaacc cctncttggt ctaatgaaat gcanttntta ntgcanagat gttntaaggt 60
gcaatatatn tnttcctttc ccgtggtttt agagccaanc tca 103

<210> 18
<211> 365
<212> DNA
<213> Homo sapiens

<400> 18
aaatgtggtc aggggtttta tagtattttt tgtttaatct ttttggttat tgaaaaaaat 60
agaacagtcc actgtccagc agaggctgct tcaactctat tgctcgcagg gctcattctg 120
catggatctg tgtttcagga tgctgcaagg acaactctgc gggcaggaag gcccccttgac 180
ccaacgctgt agcataggtc ctgctctgtg gatggggaaa gccaggggcc acatacgtcc 240
ccatgccgcc ccctccaaag actcctcgct ggtgctgagg cagggagtgg taatcttcca 300
ggttatcata ctgggacaca acagtcacac tgctctggcg cttgccgtgt ggttggtact 360

ggtac 365

<210> 19
<211> 289
<212> DNA
<213> Homo sapiens

<400> 19
ctgtaccctc cttcccctct ggcagggaga gaaggggcct cccagggact tcccctcccc 60
cttagaacag gggggtgcga gctggtatgg atgccctcct gggcttcctg ggggctctgc 120
ccactccaga ctccagtttg tccacccccc gcagggtcct acatgcctaa gagagccctt 180
gtggtagagg ccagcttgct aggcagctag gcaggagacc cctacaaggt ccaggtaaag 240
gccagggctg ccagagccag cagactggtg aggtgggacc cagcccagg 289

<210> 20
<211> 479
<212> DNA
<213> Homo sapiens

<400> 20
ctgtcctctg ttactcagat acagttccaa aactaagcga ttatataagc acatccatat 60
tttagggcta ctctaagtta aaaacctttt ctcttgtttc agagttattt acatcaaatt 120
aagacattta caaattgttc atagtataca atagcccaaa tatgattttc acctatgctg 180
tgtaaagaag ttaagcattc gtaagtttgt ctaataaatt cagtgcactt ttttccataa 240
cacgagctat tctaaatgtt ttacatttct ttcagtgcat atttccaaat tcattaaaca 300
gaatgaaatc aatgttatta aatggctata tcataatatt caagcatatt atggaatcta 360
taccacagtg ggattcacgt caatactata attcactcta gaaaaacatc acaggcacac 420
acaaaataaa gaacaaaatt tgatttttt ttataaatgt aaagtatact atctacttt 479

<210> 21
<211> 343
<212> DNA
<213> Homo sapiens

<400> 21
aaatttttta ggttaatttt cttgctgtga tatatatgag gaatttacta ctttatgtcc 60
tgctctctaa actacatcct gaactcgacg tcctgaggta taatacaaca gagcactttt 120
tgaggcaatt gaaaaaccaa cctacactct tcggtgctta gagagatctg ctgtctccca 180
aataagcttt tgtatctgcc agtgaattta ctgtactcca aatgattgct ttcttttctg 240
gtgatatctg tgcttctcat aattactgaa agctgcaata ttttagtaat accttcggga 300
tcactgtccc ccatcttccg tgttagagca aagtgaagag ttt 343

<210> 22
<211> 599
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(599)
<223> n = A,T,C or G

<400> 22
ccattgctca acttgaatgg ctgcctgggt cgggcagaag gccaggtcct catggcttcc 60
catccctaat gaccggaata catgggctgc caggtcagat gtgggccaca tgggaagtcc 120
cagctctatt ctagaaaatg catgtaccat cagcttactg atagacattt actgaacttg 180
ggtatgccag atccacaggg ggccccagag atgaggggga taagaaggtt tctgaaggca 240
tggtacagaa ggtgccagca gaggtatggg ctaggggagg cagggagagc acagagcagg 300

catcctaaag gaggcagcat ttgtgttgga gcttgaagaa gtggattgtt tgcaccgcct 360
gggcaaaggg aaggtgtgtg ttcagggcat cgagagtact gcacaaaggc tgaagcccan 420
ggcagtagag aagagaatcc actaagaagg agccaatgaa gaaaaaaaag agaaaagaat 480
cgaaggtgga tagggaaaga catgctgtcc cgcagaggtt aaaggggttc tcacctcaag 540
ccagcagttc tcaaaccttg tcagcagtgg agccctttgt tctgatgaca gcctactca 599

<210> 23
<211> 153
<212> DNA
<213> Homo sapiens

<400> 23
aaaaaggttt atgtgtgtcg aggcagttgt aaaggattta ctgcagaatc aagcccactt 60
ttaggcttag gaccaggttc taactatcta aaaatattga ctgataacaa aaagtgttct 120
aaatgtggct attctgatcc atagttgttt ttt 153

<210> 24
<211> 555
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(555)
<223> n = A,T,C or G

<400> 24
aaaggaancc ancaccatnt cagagtacat aagtggctat cagagaagcc agccgatatg 60
gattggcctg cacgacccac agaagangca gcannggcag tggattgatg gggccatgta 120
tctgtacaga tcctggtntg gcaagtccat gggtgggaac aagcactgtg ctgagatgag 180
ctccaataac aacctttaa cttggagcat gcaacgaatg caacaagcgc caacacttcc 240
tgtgcaagta ccgaccatag agcaagaatc aagattctgc taactcctgc acagncccgt 300
cctnttcctt tctgctagcc tggctaaatc tgctcattat ttcagagggg aaacctanca 360
aactaagagt gataagggcc ctactacact ggctttttta ggcttagaga cagaaacttt 420
agcattggcc cagtagtggc ttntagctct aaatgtttgc cccgccatcc ctttccacag 480
tatccttctt ccctcctccc ctgtctctgg ctgtctcgag cagtctagaa gagtgcatct 540
ncagcctatg aaaca 555

<210> 25
<211> 271
<212> DNA
<213> Homo sapiens

<400> 25
cctaggagag ggcgggggct gctgtgatcc gagagctccc tgacgcccca accttccccg 60
aacgcagcta acgagctcgt gacatccgct gacatcgcca ccggtctgct ttggagggat 120
ctagcgagag tcacctaccc cacccctact gccaggggag gggtcgttgc ccccacgagg 180
gagagaaaaa caaggactat aatgcacttc gcaaaatgta aggggccggc ttcacgccag 240
cggggccttc tgggactttg aattcaacca g 271

<210> 26
<211> 210
<212> DNA
<213> Homo sapiens

<400> 26
aaaatgggct tttgctttc taggtcatta acgtttttta tttagtttct ttagccaata 60
gtggctgagt ttcgcacttg attttcaata ttttatagta agaaatgaca aactgctttg 120

gttcatttca taaacaaact ctgcatttag ataactatta aaggttgtta agatgaagat 180
ttactgtttc tttgttactc gttggtacag 210

<210> 27
<211> 282
<212> DNA
<213> Homo sapiens

<400> 27
ctgcgtgaag atccacaacc agctcatctc gtccgtctcc aacatcacct gccccaactt 60
tgatgccagc atttgcatcc cgggctccat cacattcatg cccaatggat gctgcaaggc 120
ctgcaccccc cgcaatgaga ccagggtgcc ctgctccacc gtccccgtca ccacggaggt 180
ttcgtacgcc ggctgcacca agaccgtcct catgaatcat tgctccgggt cctgcgggac 240
atttgtcatg tactcggcca gggcccaggc cctggaccac ag 282

<210> 28
<211> 333
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(333)
<223> n = A,T,C or G

<400> 28
gtgtcggcag ggttgacctc cgtggcgagg taggtgccgt cttccacgca gtgggtaacg 60
ggcttctggc tgcacctctt gggttggcag atgatgccac ttccaccctc caggcagaca 120
cagttcttgc agtcgaactc gaagtgctcc ccaaactctc tgggcacatt gtcaggtccc 180
acacagccgc aggtcttcac gcagacatca aagccaggag cgtagttcat ggtgccctca 240
ggacagaagc agccttccac caggactgtg ttgttctgct gggaggagct gggatttgca 300
cgtgggctct tntgcanggg ccacaggacc tcg 333

<210> 29
<211> 220
<212> DNA
<213> Homo sapiens

<400> 29
aaatgtctgc atgcagccag ccatcaaata gtgaatggtc tctctttggc tggaattaca 60
aaactcagag aaatgtgtca tcaggagaac atcataaccc atgaaggata aaagccccaa 120
atggtggtaa ctgataatag cactaatgct ttaagatttg gtcacactct cacctaggtg 180
agcgcattga gccagtggtg ctaaatgcta catactccaa 220

<210> 30
<211> 435
<212> DNA
<213> Homo sapiens

<400> 30
ccagaggaga tctgcagagg ggctgcaagt tctggtctca gggtgggtaa agggtcaaag 60
aggtggctct agggcagagc tgtgtgggac caagggcttt gctgacaaca gcctcaactc 120
cagacctctc tgtggtctgt ttctcctgcc aggtccctgt tgtgcccagt gccatgcctt 180
agatggaatt gagtgtgcca gtcctaggac ctttctacga gaaaataaac cttgtataaa 240
gtataccgga cactacttca taaccacttt actctactcc ttcttcctgg gatgttttgg 300
tgtggatcga ttctgtttgg gacacactgg cactgcagta gggaagctgt tgacgcttgg 360
aggacttggg atttggtggt ttgttgacct tattttgcta attactggag ggctgatgcc 420
aagtgatggc agcaa 435

<210> 31
<211> 400
<212> DNA
<213> Homo sapiens

<400> 31
ccagccacgc ttacgttccc atcacactga tgactccggg tttggcgagc acaggagcgc 60
aaacctttc acattctttc tgtgatccaa atttgttttc gtttccacca caacctccat 120
accagaatct tgcacagctt ttggtgtttg gatcatagta ccattttaat atgaaatccc 180
tgcaagttcc ttcgtctttc ggcaacttgc atatatctgt ttcagtgaga gccaatggtt 240
ctgtgctcac cattagattg atggttgaac tagaagctga ccttgctggc tgtggaggtg 300
ggggctgaga tttctttgta ctgaaacttc cgtggtaggt ggctctgacc tgagacctca 360
ggtagcagac cacagccaca tggtatgtct gcccagcgag 400

<210> 32
<211> 325
<212> DNA
<213> Homo sapiens

<400> 32
ctgcagtttt tgactcgtcc tgggaaactg gcactgagac tcaggggtgt aacatttcac 60
ctccctgaaa tcaaatccag aaatctcagg cgagcagcgt atataaaaag ccacagggga 120
aaaaaggaat acggactcag caactcttag gtgctctgag cctcttccca agccttctgg 180
ttctgtgagc atttcattga aagaaaatgg aataacagag ttttagaaga aaaactgtat 240
ttggtcttgc aagagaaaag tatattcata taaaactcag ttctcaacta tttgccaagg 300
ttacttcttt tgtttccaac attta 325

<210> 33
<211> 292
<212> DNA
<213> Homo sapiens

<400> 33
gcttctagga ggtggcacgg tgcacgccaa gatggctgtg tccacagagg agctggaggc 60
cacggttcag gaagtcctgg ggagactgaa gagccaccag tttttccagt ccacatggga 120
cactgttgcc ttcattgttt tcctcacctt catgggcacc gtgctgctcc tgctgctgct 180
ggtcgtcgcc cactgctgct gctgcagctc ccccgggccc cgcagggaaa gccccaggaa 240
ggaaagaccc aagggagtgg ataacttggc cctggaaccc tgaccctgtg tc 292

<210> 34
<211> 112
<212> DNA
<213> Homo sapiens

<400> 34
ccaatgacac tgaccactac tttctgcgct atgctgtgct gccgcgggag gtggtctgca 60
ccgaaaacct caccccctgg aagaagctct tgccctgtag ttccaaggca gg 112

<210> 35
<211> 556
<212> DNA
<213> Homo sapiens

<400> 35
aaaatcccta ttgcaagcct aacactgacc tcgctagtaa ctcttaaggc aatcaagacg 60
gaacatgtgt ttggccccca gatgcacgaa tcctgccctc ccctcaacct ttgttcatcc 120
taacagacca acctggctcc tgcattaata tggagtgggg agaacagcaa aacaattcac 180

tgtatgtaca aaagacaatt cagtgcaaac ctagaaactt ctcttagtca atagtttcca 240
attttctgag acgaggtctt gctccatcac ccaggccgga gtgcagtggc acgatcttag 300
ctcactgcaa actccacctc ccaggctcac gggatcctcc cacctcagcc tcccgagtag 360
ctgggactac aggcatgcac caccacaccc agctaatttt tgtattttta gtagagacag 420
ggttttgcca tgttgcccag gctggtctcg aactcctggg ctcaagcaat tcgcccacct 480
gagcctccca aagtgctggg attacaggcg tgaaccacca cacccggcct caaattttct 540
tattgccctt gagcaa 556

<210> 36
<211> 404
<212> DNA
<213> Homo sapiens

<400> 36
ttggttactt cttttaatgt attcaaaaat gttgaacaca tacagaactg aattaagaag 60
caacaactgc cctatggaag agctgtatta gtacagaatg cttttaagaa ccaaggacaa 120
attttcagta ttgaagaaga caacatacat aaaaagcact ccaaattcat ttctaattcc 180
ttcaatacca tgctaaagtt ctttttaga gggtatgtct cttaacaact ttacataatt 240
cacaatgaga atgtgacaac atgtcaattt ggcaatcaac acttcttcat tgcaccttac 300
ttactttatg catgcggccc acacattact tcagctcaag aggctggggt aattctgtcc 360
ctaaggcaat caaggagctc agcacaaacc ttgaaatcat tttt 404

<210> 37
<211> 344
<212> DNA
<213> Homo sapiens

<400> 37
tgccaggttc acacatccca ggaaaaaaga agcataaaaa gcattagcag tcagtgactg 60
atgataatgc tgcaataatg ggaatggttt tgtttctaaa ccaaattatt tctaaatcaa 120
atcatttatt gctttgtttc taaagcaatt gagtcactaa gtttgtgaac tgtaggagaa 180
cacatcaaga ttgaatcctg tgttaagcag aaggtaaaac cagagccagg cgcagtggct 240
cgtgcctgta attccaaaac cttggcagga agatcgattg aggccaggag ctcaagacga 300
gcctgggcaa catagaaaga ccctatcttt acaaaaaaaa cttt 344

<210> 38
<211> 343
<212> DNA
<213> Homo sapiens

<400> 38
cagaaagctg tggatgagat gaacggaaag gagctcaatg gaaaacaaat ttatgttggt 60
cgagctcaga aaaaggtgga acggcagacg gaacttaagc gcaaatttga acagatgaaa 120
caagatagga tcaccagata ccagggtgtt aatctttatg tgaaaaatct tgatgatggt 180
attgatgatg aacgtctccg gaaagagttt tctccatttg gtacaatcac tagtgcaaag 240
gttatgatgg agggtggtcg cagcaaaggg tttggttttg tatgtttctc ctccccagaa 300
gaagccacta aagcagttac agaaatgaac ggtagaattg tgg 343

<210> 39
<211> 272
<212> DNA
<213> Homo sapiens

<400> 39
cattacataa aaggacatac ctctacctag caatgaccat actgcatgaa gagggactaa 60
gatgaaggaa agaaacaaaa aggcaagtca aagaaagaca ggagtgtaag gtcctaagga 120
aggggcaaat aatcaaaggt ctcatctgat caggagcaat gccaatccaa tccagttcta 180
ggtcccaagt aaagaaacat agtctgagaa aagaggccag ggatacagct tgggatgttc 240

agagagtgga aatgacagag gtggattatt tt 272

<210> 40
<211> 414
<212> DNA
<213> Homo sapiens

<400> 40
ctggtccagc agctcgagca gtgggagttc cgagtctgtg gagtttttct tgttgattct 60
cagttcatgg tggagtcatt caagtttatt tctggcatct tggcagccct gagtgccatg 120
atctctctag aaattccgca agtcaacatc atgacaaaaa tggatctgct gagtaaaaaa 180
gcaaaaaagg aaattgagaa atttttagat ccagacatgt attctttatt agaagattct 240
acaagtgact taagaagcaa aaaattcaag aaactgacta aagctatatg tggactgatt 300
gatgactaca gcatggttcg atttttacct tacgatcagt cagatgaaga aagcatgaac 360
attgtattgc agcatattga ttttgccatt caatatggag aagacctaga attt 414

<210> 41
<211> 174
<212> DNA
<213> Homo sapiens

<400> 41
cctacgagaa aatccttttc actgaggcca cccggatcct cttcttcaac acacccaaaa 60
agatgacaga ctacgccaag aagcgagggt gggtcctggg ccccaacaac tactacagtt 120
ttgccagcca gcagcagaag ccggaagaca ccaccattcc ctccacagaa ctgg 174

<210> 42
<211> 260
<212> DNA
<213> Homo sapiens

<400> 42
ctggtggtga gctagctcat ctccacaaca cccaccaagg ctccaaagac acctctcagc 60
ttagctttca ctggtggatt ggatgctgtt ctcatttgat ctatagatcc tagcattttg 120
gagtttgatg aaagaagggc aggagagtat taagaatcaa aggtcttagc tgggcttggt 180
ggctcatgcc tgtaattcca gcactttggg aggccgagga gggtggatca cttgaggcca 240
ggagtttgag accaccctgg 260

<210> 43
<211> 566
<212> DNA
<213> Homo sapiens

<400> 43
cctctgtgca agcagcacat aggatctgga tgtaggttga ggatagatcc tcacccacca 60
gtggggtaac tttcccagca attctgaaac taaaataagg aaggcacatt cccagagccc 120
tgctgagtag gggcttcagg ctattttcac tctacacaaa atgggggaga ggagttccct 180
ctccactaat ttttcaccca taaacctcca catcactagg aacctaaagg ggaactccaa 240
aggccaacac atccttggtg gttatatgtg ttgtcctgac aacctcctgc tccagaaatg 300
ccaggagcat tggatatgtc attgggagca tcaggcagtc caacatcgga gggagaaagg 360
cccagagatg aggatctgag tcaggctggc aaggctggag tcagaaagtg accattaggc 420
aactggtcac tacaattggt ggctacaaag aagtggtcac agtcaccaaa ataaagaggt 480
ttacaacaac ggtttcccct taggtcattt tgaccaggac agtaccctaa aggaaataag 540
gcagcatcgc ataaagcaag agcccc 566

<210> 44
<211> 344
<212> DNA

<213> Homo sapiens

<400> 44
ccactggctg agttattggc ctggcaggta tagagtccgc tgttcttctc agtgatgttg 60
gagataaaga gctcttgtgt gtgttgctgg atgttcccat caatcagcca agaatactgt 120
gcaggtgggt tagaggctgc atggcaggag aggctgaggt tcacccctgg acggtaatag 180
gtgtatgagg gggaaatggt ggggtcgtct gggccataga ggacattcag gatgactggg 240
tcgctgtggt caacacttaa ttcgttctgg attccacact catagggtcc tacatcattc 300
cttgtgacac tgagtagagt gagggtcctg ttgtcattgg acag 344

<210> 45
<211> 404
<212> DNA
<213> Homo sapiens

<400> 45
ttggttactt cttttaatgt attcaaaaat gttgaacaca tacagaactg aattaagaag 60
caacaactgc cctatggaag agctgtatta gtacagaatg cttttaagaa ccaaggacaa 120
attttcagta ttgaagaaga caacatacat aaaaagcact ccaaattcat ttctaattcc 180
ttcaatacca tgctaaagtt ctttttaga gggtatgtct cttaacaact ttacataatt 240
cacaatgaga atgtgacaac atgtcaattt ggcaatcaac acttcttcat tgcaccttac 300
ttactttatg catgcggccc acacattact tcagctcaag aggctggggt aattctgtcc 360
ctaaggcaat caaggagctc agcacaaacc ttgaaatcat tttt 404

<210> 46
<211> 215
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(215)
<223> n = A,T,C or G

<400> 46
gtgggtgaca gtgatgccag gctcgcccac tactgcactg gacacagcct caccaatgcc 60
accttcataa taatggtcct ccacggtgag gatcctgccc ttggtggcac gagcgctgtc 120
gagaatgagt tttctgtcca ggggcttgat ggtgaanngg tccagcacgc ggatgttgat 180
cttttctttc ttcagcagtt cggcagcggc caagg 215

<210> 47
<211> 425
<212> DNA
<213> Homo sapiens

<400> 47
aaattataag tattgtgaat tcacactctc aggctattgt ctgacttgat ctacgtctca 60
taaagcctgt acctgagtgg agtggaaggt ggagtcttag gttaatcagt tactgactct 120
accctcaccc tctttcaatt gaggtaaact ttgctgtttt tctttttcat aaagcattct 180
caaattgttg agtttattgc tgaaaaaaat ctccatgact ttacagatag aattacaaac 240
taaatgatgt cttgtattta gaagcagagt acagacctaa cgaactgtta gattctccac 300
catcacttag ggtttgccca gaagcaacac cagagaatta cagacaacgc gcttttgctg 360
aacaagcatt tgtagcttgt acaatggcag aatgggccaa aagcttagtg ttgtgacctg 420
ttttt 425

<210> 48
<211> 423
<212> DNA

<213> Homo sapiens

<400> 48
ctgctgcaac attaccgtct gcaagtgcaa caccagcctg tgcaaagaga agccctccgt 60
gtgcccgctg ggattcgaag tgaagagcaa gatggtgcct ggaaggtgct gtcctttcta 120
ctggtgtgag tccaaggggg tgtgtgttca cgggaatgct gagtaccagc ccggttctcc 180
agtttattcc tccaagtgcc aggactgcgt gtgcacggac aaggtggaca acaacaccct 240
gctcaacgtc atcgcctgca cccacgtgcc ctgcaacacc tcctgcagcc ctggcttcga 300
actcatggag gcccccgggg agtgctgtaa gaagtgtgaa cagacgcact gtatcatcaa 360
acggcccgac aaccagcacg tcatcctgaa gcccggggac ttcaagagcg acccgaagaa 420
caa 423

<210> 49
<211> 121
<212> DNA
<213> Homo sapiens

<400> 49
ccagggcggt acgaatcgtc tcctggcact gtgcaggccc acagctgaga actggcctct 60
acaaatccca gagaccgtgc gtaacacaca tcaagacaga acctgttgcc attttcagcc 120
a 121

<210> 50
<211> 253
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(253)
<223> n = A,T,C or G

<400> 50
ctggggcggc ctatgccgag tggcgcccat ggcgaagagg gctcagctcg catgtggaag 60
actctcacct tcttcgtcgc gctccccggg gtggcagtca nnatgctgaa tgtgtacctg 120
aagtcncacc acggagagca cgagagaccc gagttcatcg cctaccccca tctncgcatc 180
aggaccaanc cgtttccctg gggagatggt aaccatactc tattccataa ccctcatgtg 240
aatccacttc caa 253

<210> 51
<211> 228
<212> DNA
<213> Homo sapiens

<400> 51
ctgaaagtaa acagaatgga ttgccagtta catgtatgcc tgcccagttc ccttttttatt 60
tgcagaagct gtgagttttg ttcacaatta ggttcctagg agcaaaacct caaggattga 120
tttattgttt tcaactccaa ggcacactgt taataaacga gcagggtgtt ttctctcttc 180
ctttctaata tatggagttt cgaagaataa aatatgagag caatattt 228

<210> 52
<211> 217
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(217)

<223> n = A,T,C or G

<400> 52
ctgactagtt cccctaataa tcggtgcccc cgatatggcg tntccccgca taaacaacat 60
aagcttctga ctcttacctn cctctctcct actcctgctc gcatctgcta tagtggaggc 120
cggagcagga acaggttgaa cagtctaccc tcccttagca gggaactact cccaccctgg 180
agcctccgta gacctaacca tcttctcctt acaccta 217

<210> 53
<211> 186
<212> DNA
<213> Homo sapiens

<400> 53
aaattttcat tgagttgtcc atctccagca tatagggctt caggagcaga gcagaccttg 60
tttttagtgg ttccatggga taaaatggga ttggaggagc tagaagaatt cagggtctgg 120
tccaatctgc cagtcttcct gaaatatcga aaatacacca gggctgctat atcagagcca 180
ccctgg 186

<210> 54
<211> 164
<212> DNA
<213> Homo sapiens

<400> 54
caggcgcagc ccagcctcga aatgcagaac gacgccggcg agttcgtgga cctgtacgtg 60
ccgcggaaat gctccgctag caatcgcatc atcggtgcca aggaccacgc atccatccag 120
atgaacgtgg ccgaggttga caaggtcaca ggcaggttta atgg 164

<210> 55
<211> 330
<212> DNA
<213> Homo sapiens

<400> 55
ctgtgatgaa cagtacttgt gtcagttctg tgaacatgaa actaatgatc cagaagactt 60
gcatagccat gtggtaaatg agcatgcatg taaattaata gagttaagtg ataagtataa 120
caatggtgaa catggacagt atagcctctt aagcaaaatt acctttgaca aatgtaaaaa 180
cttctttgta tgtcaagtat gtggttttcg gagtagactt cacacaaatg ttaacaggca 240
tgttgctatt gaacatacaa aaatttttcc ccatgtttgt gatgactgtg ggaaaggctt 300
ttcaagtatg ctagaatatt gcaagcattt 330

<210> 56
<211> 408
<212> DNA
<213> Homo sapiens

<400> 56
cctagtatga ggagcgttat ggagtggaag tgaaatcaca tggctaggcc ggaggtcatt 60
aggagggctg agagggcccc tgttaggggt catgggctgg gttttactat atgataggca 120
tgtgattggt gggtcattat gtgttgtcgt gcaggtagag gcttactaga agtgtgaaaa 180
cgtaggcttg gattaaggcg acagcgattt ctaggatagt cagtagaatt agaattgtga 240
aaatgataag tgtagaggga aggttaatgg ttgatattgc tagggtggcg cttccaatta 300
ggtgcatgag taggtggcct gcagtaatgt tagcggttag gcgtacggcc agggctattg 360
gttgaatgag taggctgatg gtttcgataa taactagtat ggggataa 408

<210> 57
<211> 218

<212> DNA
<213> Homo sapiens

<400> 57
ccttatgaca tgtgctgtgg ctcccagcac cagtttaggt acttggagtg cagcagggaa 60
gaaaataact tggctgctct gcacgctggg ggcttcactc agcggcatct agacagacac 120
ataattggcc gggcgtggcg gctcacgcct gtaatcccaa aacctgggag gccgaggcag 180
gccgatcact tgaggtcagg agttcgagac cagcctgg 218

<210> 58
<211> 390
<212> DNA
<213> Homo sapiens

<400> 58
ccaagaacgt gcaataaatt ggaagtttgc cccggggcag caagaattta tgctgccatt 60
gaaaagcagg taccagtgcc ccttttcaga cagtttttga ttcgctctag actttttttt 120
ttttaatagg gagggaaaaa atttgataat tttctttttt ctacatgcac ttaagactaa 180
aacacaggtt tggattaatt ttatttgctt ccttttttccg cttttcttcc cgcagagcct 240
gatgggagaa tgtccagggc agggaaacca catttttgt aggtgataac tcaatgaaaa 300
ttggtgctta ttttttacac ttctctcttg tggctctctt gtggtgctat ctgttttaag 360
gtctccttga aggcgcactg gggtccctgg 390

<210> 59
<211> 516
<212> DNA
<213> Homo sapiens

<400> 59
ttgttttgct tcttccttaa agcatttgca acagctacag tctaaaattg cttctttacc 60
aaggatattt acagaaaaga ctctgaccag agatcgagac catcctagcc aacatcgtga 120
aaccccatct ctactaaaaa tacaaaaatg agctgggctt ggtggcgcac acctgtagtc 180
ccagttactc gggaggctga ggcaggagaa tcgcttgaac ccgggaggtg gagattgcag 240
tgagcccaga tcgcaccact gcactccagt ctggcaacag agcaagactc catctcaaaa 300
agaaaagaaa agaagactct gacctgtact cttgaataca agtttctgat accactgcac 360
tgtctgagaa tttccaaaac cttaatgaac taactgacag cttcatgaaa ctgtccacca 420
agatcaagca gagaaaataa ttaatttcat gggactaaat gaactaatga ggataatatt 480
ttcataattt tttatttgaa attttgctga ttcttt 516

<210> 60
<211> 222
<212> DNA
<213> Homo sapiens

<400> 60
cctcttttta ccagctccga ggtgattttc atattgaatt gcaaattcga agaagcagct 60
tcaaatctgc cggggcttct cccgcctttt ttcccggcgg cgggagaagt agattgaagc 120
cagttgatta gggtgcttag ctgttaacta agtgtttgtg ggtttaagtc ccattggtct 180
agtaagggct tagcttaatt aaagtggctg atttgcgttc ag 222

<210> 61
<211> 350
<212> DNA
<213> Homo sapiens

<400> 61
aaaaaactca aaaagctggg aattaagtgg tttcagtaat aatgctatac cgaggtgctt 60
gcattgtatt tcataatttt gttacaaacc aaaattattt ttaatgagaa cagtcttggg 120

ttcagaggtg tgatgccaga atgtattttc gtactgttag gcccttggaa cagataccgg 180
tgctttctga aagatgaaag aaatgcaatg ggtgctcttc atgcaaggtt gcaaacctac 240
caagaatgca taatagtctc acttttcccc aataaagaga tgcgtgtgac tagttttgga 300
cttttaacct taatgggggt tgcatgtctc ctattgttaa tcattgtcag 350

<210> 62
<211> 391
<212> DNA
<213> Homo sapiens

<400> 62
aaaaaccaga tcgctaccca tgagaagaaa gctcatgaaa actggctcaa agctcgtgct 60
gcagaaagag ctatagctga agagaaaagg gaagctgcca atttgagaca caaattatta 120
gaattaacac aaaagatggc aatgctgcaa gaagaacctg tgattgtaaa accaatgcca 180
ggaaaaccaa atacacaaaa ccctccacgg agaggtcctc tgagccagaa tggctctttt 240
ggcccatccc ctgtgagtgg tggagaatgc tcccctccat tgacagtgga gccacccgtg 300
agacctctct ctgctactct caatcgaaga gatatgccta gaagtgaatt tggatcagtg 360
gacgggcctc tacctcatcc tcgatggtca g 391

<210> 63
<211> 439
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(439)
<223> n = A,T,C or G

<400> 63
aaaataggcc ctgagtataa gagcatgaag agctgccttt atgtcggcat ggcgagcgac 60
aacgtcgatg ctgctgagct cgcggagacc attgcggcca cagcccggga gatagaggag 120
aactcgaggc ttctggaaaa catgacagaa gtggttcgga aaggcattca ggaagctcaa 180
gtggagctgc agaaggcaag tgaagaacgg cttctggaag agggggtgtt gcggcagatc 240
cctgtagtgg gctccgtgct gaattggttt tctccggtcc aggctttaca gaagggaaga 300
acttttnaac ttgacagcag gctctctgga gtccacagaa cccatatatg tctacaaagc 360
acaaggtgca ggagtcacgc tgcctccaac gccctcgggc agtcgcacca agcagaggct 420
tccaggccag aagcctttt 439

<210> 64
<211> 249
<212> DNA
<213> Homo sapiens

<400> 64
aaaacatttt ttagtctgta atacactcca cttgaagcac ttaagtcttc cttaaatgac 60
ttttcttaag taatgatact gtgtgttttc ccaaagcaca cagtatcatt acttaagaaa 120
atttttataa attactatct gttgaaaagg tgtccttttc ctttcttcta gtatttttt 180
cttaccaaaa ttcactaatc ttgaatgttt gtgatattaa atttcaaatg cagaatactt 240
gactcattt 249

<210> 65
<211> 229
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature

<222> (1)...(229)
<223> n = A,T,C or G

<400> 65
ggagctcang cggtgatgtt cgctcacctg ctgctcacct actgctgcgt ggcccagttn 60
ctaacagacc acagacggat ctgctgggga ctcctgcata taaagtctgc catcatggat 120
gtnntcgcag aagcaaatgg caccttgcc ttaaaccttt tgaaaacgct gggtaaagac 180
aactcgaaga atgtgnnttt ctcacccatg agcatgtnct gtgccctgg 229

<210> 66
<211> 195
<212> DNA
<213> Homo sapiens

<400> 66
ccacagaccc ccaggtcatt gtgttcactg tactctgtgg gcaaggatgg gtccagaaga 60
ccccacttca ggcactaaga ggggctggac ctggcggcag gaagccaaag agactgggcc 120
taggccagga gttcccaaat gtgaggggcg agaaacaaga caagctcctc ccttgagaat 180
tccctgtgga ttttt 195

<210> 67
<211> 425
<212> DNA
<213> Homo sapiens

<400> 67
ctgtcaacct tgacaaattg tggactttgg tcagtgaaca gacacgggtg aatgctgcta 60
aaaacaagac tggggctgct cccatcattg atgtggtgcg atcgggctac tataaagttc 120
tgggaaaggg aaagctccca aagcagcctg tcatcgtgaa ggccaaattc ttcagcagaa 180
gagctgagga gaagattaag agtgttgggg gggcctgtgt cctggtggct tgaagccaca 240
tggagggagt ttcattaaat gctaactact ttttccttgt ggtgtgagtg taggttcttc 300
agtggcacct ctacatcctg tgtgcattgg gagcccaggt tctagtactt agggtatgaa 360
gacatggggt cctctcctga cttccctcaa atatatggta aacgtaagac caacacagac 420
gttgg 425

<210> 68
<211> 471
<212> DNA
<213> Homo sapiens

<400> 68
ctgtgtgact gcctgtccct acaactacct ttctacggac gtgggatcct gcaccctcgt 60
ctgccccctg cacaaccaag aggtgacagc agaggatgga acacagcggt gtgagaagtg 120
cagcaagccc tgtgcccgag tgtgctatgg tctgggcatg gagcacttgc gagaggtgag 180
ggcagttacc agtgccaata tccaggagtt tgctggctgc aagaagatct ttgggagcct 240
ggcatttctg ccggagagct ttgatgggga cccagcctcc aacactgccc cgctccagcc 300
agagcagctc caagtgtttg agactctgga agagatcaca ggttacctat acatctcagc 360
atggccggac agcctgcctg acctcagcgt cttccagaac ctgcaagtaa tccggggacg 420
aattctgcac aatggcgcct actcgctgac cctgcaaggg ctgggcatca g 471

<210> 69
<211> 352
<212> DNA
<213> Homo sapiens

<400> 69
gtgtccttta tcttacttta tctgtacagt aatcctgtga gaaagacagg acagaaacca 60
ctgtgcctat tttacagata cgaaaactga gacacaggta aggggcttgt ctgtagtccc 120

atagctagca gatggctgga gccaagactg aggctcgttc ttcaatgctg agccagggct 180
ccttccgctg caccacaaga acgctagacc actcgccacc agccttctca ttccctcttc 240
ctccattcta atcatttcta gctggctggc ctccacagag cataggaaaa cagccagggc 300
cgggcacggt ggctcatgcc tgtaatctca acactctggg aggccaaggc ag 352

<210> 70
<211> 519
<212> DNA
<213> Homo sapiens

<400> 70
aaaaaaagct atgtcttcac tccaaaatga cagagacaga ctactgaagg aattgaagaa 60
tctgcagcag caacacttac agattaatca agagatcact gagttacatc cactgaaggc 120
tcaacttcag gagtatcaag ataagacaaa agcatttcag attatgcaag aagagctcag 180
gcaggaaaac ctctcctggc agcatgagct gcatcagctc aggatggaga agagttcctg 240
ggaaatacat gagaggagaa tgaaggaaca gtaccttatg gctatctcag ataaagatca 300
gcagctcagt catctgcaga atcttataag ggaattgagg tcttcttcct cccagactca 360
gcctctcaaa gtgcaatacc aaagacaggc atccccagag acatcagctt ccccagatgg 420
gtcacaaaat ctggtttatg agacagaact tctcaggacc cagctcaatg acagcttaaa 480
ggaaattcac caaaaggagt taagaattca gcaaacctc 519

<210> 71
<211> 434
<212> DNA
<213> Homo sapiens

<400> 71
ctgtatgtga taatgaaagg gtttttcttt cttatgttaa atacaagcga agtgattaac 60
tggaagatag cgtctgattg cgaggaaatc agtgattcag atggtgtggg aatggcacct 120
ggggatgggg gaggcaggac ggagatggag gaagctggtg cagcctagcc tgccttgtgc 180
caaggacacc caagggcaga gggactgagc tctgggggag gacagatttg acataactgg 240
tccagcctca cagtttacag gtcctggagg gtgaggaaca gacgtgggag caccagaggg 300
acagagctga tggcctgacg ctctcttcag gagggcaccc ccaaggggcc tctgcttcct 360
cagtgccccc tgagctttat cagcagaggg gtgttttcca gccacaagga gctgtatcta 420
acactaatgc cttt 434

<210> 72
<211> 295
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(295)
<223> n = A,T,C or G

<400> 72
ccattctagt gatccaaagg ccgtaatgtt ccccacctac aaatatgttg acatcaacac 60
atttcnnctc tctgctgatg acatanntgg cattcagtcc ctgtatggag acccaaaaga 120
gaaccaacgc ttgccaaatc ctgacaattc agaaccagct ctctgtgacc ccaatttgag 180
ttttgatgct gtcactaccg tgggaaataa gatcttttc ttcaaagaca ggttcttctg 240
gctgaaggtt tctgagagac caaagaccag tgttaattta atttcttcct tatgg 295

<210> 73
<211> 118
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(118)
<223> n = A,T,C or G

<400> 73
ctgctgtctg acnatgaaac caaagacgac atgnncatgt cctcctactc ggtggtcagc 60
acgggctncc tgcaatgtga agaccttgca nacnacacgg tgctggtggg cggggagg 118

<210> 74
<211> 633
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(633)
<223> n = A,T,C or G

<400> 74
ccttggtcgc ccaacagccc attggcaacc ttctcatgtc tggtcacaga gaactgattt 60
gtcctcagta cctctccgtc caccttcatg tacacagtgg gcaccacctt cacaaagtac 120
tggaacatca tggaggcttg gggcgcagtg acattggtgt ggtccagggg gttcacaatg 180
cctggatagt cctccccaaa tgacaggtgc tggatgtagt gggtcatgtt gatgttgtca 240
aggccaaagc tctgcaagtc atggacgtgc acatgggact gctggaagct cttcccaggg 300
gcaaagtgga agtttccggc caccttattg acttccaaga agccatacac ctggcagcct 360
tcattcttct gctcctgcat cttctggctg aagccctctc gccggcactg ctcaatagta 420
tctgggttct tgaaggccca gcctctacgg cgatatgcct cccgcacatc ttcacaggtg 480
ttacagcact tgatatcttc tgcctcagca ccatagcagc tctcacagcg atcagggtcc 540
agggagtcag ggtcaaacac cgtcacctcg actttnccca agctcatgcc gntcagcctc 600
tgagctcacg gggatgccat ctttatctag tcg 633

<210> 75
<211> 305
<212> DNA
<213> Homo sapiens

<400> 75
ttgccaaagc ctccgattat gatgggtatt actatgaaga agattattac aaatgcatgg 60
gctgtgacga taacgttgta gatgtggtcg ttacctagaa ggttgcctgg ctggcccagc 120
tcggctcgaa taaggaggct tagagctgtg cctaggactc cagctcatgc gccgaataat 180
aggtatagtg ttccaatgtc tttgtggttt gtagagaata gtcaacggtc ggcgaacatc 240
agtgggggtg aggtaaaatg gctgagtgaa gcattggact gtaaatctaa aagacagggg 300
ttagg 305

<210> 76
<211> 611
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(611)
<223> n = A,T,C or G

<400> 76
ctgtttcata ggctggagat gcactcttct agactgctcg agacagccag agacagggga 60
ggagggaaga aggatactgt ggaaagggat ggcggggcaa acatttagag ctagaagcca 120

ctactgggcc aatgctaaag tttctgtctc taagcctaaa aaagccagtg tagtagggcc 180
cttatcactc ttagtttgct aggtttcccc tctgaaataa tgagcagatt tagccaggct 240
agcagaaagg aagaggacgg ggctgtgcag gagttagcag aatcttgatt cttgctctat 300
ggtcggtact tgcacaggaa gtgttggcgc ttgttgcatt cgttgctgct ccaagttaaa 360
aagttgttat tggagctcat ctcagcacag tgcttgttcc cacccatgga cttgccagac 420
caggatctgt acagatacat ggccccatca atccactgcc actgctgcct cttctgtggg 480
tcgtgcaggc caatccatat cggctggctt ctctgatagc cacttatgta ctctgctatg 540
gtgctggctt cctttaaact cangatagat gccaggtggg ctccgtttcc gtaagactga 600
cactcgagct c 611

<210> 77
<211> 267
<212> DNA
<213> Homo sapiens

<400> 77
ctggtatcag agaagtcagt agaggtcact gagaccggca gtctttcttg ctttttgcat 60
tagtgccctc aggacacaca gcaaacagtg atcatgagaa gagtgagctc aatagttttc 120
catcaagtgt gcttaaaatt ccatgcagtc gccataaggg tacaacttct gaggtatggt 180
caacctatgg tacattagta aatgataagg ggaggaagaa atgaaaacct aaacgtctac 240
tgcaatgaaa accaatagca atgtcag 267

<210> 78
<211> 295
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(295)
<223> n = A,T,C or G

<400> 78
aaatatttat cagtctaaac ttgtgcagtg tagtaaacat gcaagttgtt acgattgagc 60
tgtattacca taagtagaat tttaagtaaa ctggtgaatt tgggcaataa atgtttttgc 120
tttttgtttg attttttttt acaagctaac tgttagaggt atacatttat ttatctgttg 180
tacagatttg attatgattt taatgtttga aagattgcac ttgtttgctt ttactatatg 240
tggggtaaaa tatattttnt gntcacagta tatgaaaata tggagtaatt tacct 295

<210> 79
<211> 320
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(320)
<223> n = A,T,C or G

<400> 79
ttttttttnt ttttttttnt tttggggana acagggtgtc gctatattac ccagtcaggt 60
ctcgaactct ggacctcaag tgacccacct gcctcggcct cccaaggtgc tgggatagca 120
ggcgtgagcc actgtgccca gcctcaccta atggtttctt agcaaacttc agtanaatgt 180
ttanaacgcg gccctgataa acttgagtgc tggtaggagg tgctacctcg ctcaatctgt 240
gagcaaccag ccctgtgccc tggatgcttg gcgggtggag agaaanacag tgttatgtgg 300
gcaagcctcc aaactcacca 320

<210> 80

<211> 133
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(133)
<223> n = A,T,C or G

<400> 80
tgagggtctt actcttttag tataaatagt accgntaact tccaattaac tagttttgac 60
aacattcaaa aaagagtaat aaacttcgcc ttaattttaa taatcaacac cctcctagcc 120
ttactactaa taa 133

<210> 81
<211> 406
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(406)
<223> n = A,T,C or G

<400> 81
ctgtgggggc ctcctttcct agtttctgaa tgatcttcct gtggctctgt gagcaggccc 60
agcatgggga atgggctaaa aggcttatac atctcttttg gccctcagat gcacttaccc 120
ttttctttgg tgccctcttt ccccaagaga atattcaggc caattttgct tttttccttg 180
tttctgcatt agtaagacat tataaactag caacttgtaa tacctctaac tctcactgtc 240
ttatgttagt ataaagtacc tcaaggtaat aagaatgtgg aacttaaatg ccacttacag 300
aaagtcaaac aaagcccatg tcacactttg atgaatncaa agtattaaat cttancaact 360
gatgaagtaa aaagctattt ttgctaangt ttaactattg gacttt 406

<210> 82
<211> 340
<212> DNA
<213> Homo sapiens

<400> 82
aaaaattatg agccttttct agcccccacc ttcccaaccc tcagagaagg acagtaaaga 60
aatcagtgga tccaggtatt tacctgttgt tgaattgtga ggttgtgagg tagacgtgta 120
acaaggacaa ggaagtttgg ggatctgctt ggagaatgaa ggtttattca aaacaagtgg 180
acaggtcagg ggtaacgggt gatgagggca cctggctttt gtaatcatgt ggggactgtc 240
ccctgggagg tgcagcaaaa atcagaacgg agacagaagc tcacagtctg gttttagctg 300
ccaactccta tggaagtcca tagctgactc ctatggaagg 340

<210> 83
<211> 380
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(380)
<223> n = A,T,C or G

<400> 83
gtgcgcacca ccacacccgg ccaattttgg tntcttttgt aatgacanaa tnncgccatg 60

ttgcccangc tgatctcaaa ctcncgggct caagcaatcc acccacttcg gcctaccaaa 120
gtgctaggat cacaggcatg agccatagca cctggcccac attttctttt gttaaatgaa 180
gttaatctat gtnctagtaa atanacaatt atgtttccaa cacaacagaa atctatttca 240
acactaaaca tcactgaacg attttgctaa ggttttcatg ctagatgtgt cttactaaca 300
aaggtaacac aattccacag ttctgttttt gaataaaagg natatatgtt atatatctga 360
aaacttacac aagatgttca 380

<210> 84
<211> 529
<212> DNA
<213> Homo sapiens

<400> 84
aaaataattt agttttgctt gcttccattg atcagtcttt tacttgaggc attaaatatc 60
taattaaatc gtgaaatggc agtatagtcc atgatatcta aggagttggc aagcttaaca 120
aaacccattt tttataaatg tccatcctcc tgcatttgtt gataccacta acaaaatgct 180
ttgtaacaga cttgcggtta attatgcaaa tgatagtttg tgataattgg tccagtttta 240
cgaacaacag atttctaaat tagagaggtt aacaagacag atgattacta tgcctcatgt 300
gctgtgtgct ctttgaaagg aatgacagca gactacaaag caaataagat atactgagcc 360
tcaacagatt gcctgctcct cagagtctct cctatttttg tattacccag ctttcttttt 420
aatacaaatg ttatttatag tttacaatga atgcactgca taaaaacttt gtagcttcat 480
tattgtaaaa catattcaag atcctacagt aagagtgaaa cattcacaa 529

<210> 85
<211> 525
<212> DNA
<213> Homo sapiens

<400> 85
aaatccagaa gaaaacaaaa aattctggac ttggattttt gaggttagaa agaatttttg 60
aagcagaaat cccaataaat atgctgatct tcctaagaat gaataaatca cagaatttta 120
agccaaaagg aaagggcaca taaagattat ccagtccaac ctcatttac agacagggag 180
ggtgacctgc ccaaagtcac atgactaaaa ggagaagggg tggccttgga atacagatgt 240
tctgacttct agggtcttat cttttaatgt tgccttttg tcctcaaagc tgcctgctta 300
ttgggttgga agaactcaca tcttatgaag ggttagaccc tgccttgaaa atcagtatgg 360
taggctgggc gtgggccctt acgcctgtaa tcctatcgct ttgggaggcc cagccaggtg 420
gattgcttga gcccagcagc ttcagaccag cctgggcaac acggcaaaac cccatctcta 480
caaaaaaatt caaaaattag ctgggcatgg tggtgcacac acctc 525

<210> 86
<211> 430
<212> DNA
<213> Homo sapiens

<400> 86
aaaaaatatt tagctttgca gttcctgacc ccttaatgcc tgacccttcc aagcaaccaa 60
agaaccagct taatcctatt ggttcattac aggaattggc tattcatcat ggctggagac 120
ttcctgaata tacccttcc caggagggag gacctgctca taagagagaa tatactacaa 180
tttgcaggct agagtcattt atggaaactg gaaaggggc atcaaaaaag caagccaaaa 240
ggaatgctgc tgagaaattt cttgccaaat ttagtaatat ttctccagag aaccacattt 300
ctttaacaaa tgtagtagga cattctttag gatgtacttg gcattccttg aggaattctc 360
ctggtgaaaa gatcaactta ctgaaaagaa gcctccttag tattccaaat acagattaca 420
tccagacctc 430

<210> 87
<211> 408
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(408)
<223> n = A,T,C or G

<400> 87
ccatgtacat atgggtcctc gaagacaagc catgaaagag atgtccatcg atcaagccaa 60
atatcagcga tggcttatta agaacaaaat gaaggcattt tatgctccag tacatgcaga 120
tgacttgaga gaaggtgcac agtatttgat gcaggctgct ggtcttggtc gtatgaagcc 180
aaacacactt gtccttggat ttaagaaaga ttggttgcaa gcagatatga gggatgtgga 240
tatgtatata aacttatttc atgatgcttt tgacatacaa tatggagtag tggttattcg 300
cctaaaagaa ggtctggata tatctcatct tcaaggacaa gaagaattat tgtcatcaca 360
agagaaatct cctggcacca aangatgtgg tagtnagtgt ggaatata 408

<210> 88
<211> 502
<212> DNA
<213> Homo sapiens

<400> 88
aaaaaagttt ccacttgaca ctttgatccc tgatggaaaa cgcataatct gggacagtag 60
aaagggcttc atcatatcaa atgcaacgta caaagaaata gggcttctga cctgtgaagc 120
aacagtcaat gggcatttgt ataagacaaa ctatctcaca catcgacaaa ccaatacaat 180
catagatgtc caaataagca caccacgccc agtcaaatta cttagaggcc atactcttgt 240
cctcaattgt actgctacca ctcccttgaa cacgagagtt caaatgacct ggagttaccc 300
tgatgaaaaa aataagagag cttccgtaag gcgacgaatt gaccaaagca attcccatgc 360
caacatattc tacagtgttc ttactattga caaaatgcag aacaaagaca aaggacttta 420
tacttgtcgt gtaaggagtg gaccatcatt caaatctgtt aacacctcag tgcatatata 480
tgataaagca ttcatcactg tg 502

<210> 89
<211> 329
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(329)
<223> n = A,T,C or G

<400> 89
ttgtgatcgt ggtgtgcgtc agcttcctgg tgttcatgat tatcctgggg gtatttcgga 60
tccgggccgc acatcggcgg accatgcggg atcaggacac cgggaaggag aacgagatgg 120
actgggacga ctctgccctg accatcaccg tcaaccccat ggagacctat gaggaccagc 180
acagcagtga ggaggaggag gaagaggaag aggaagagga aagcgaggac ggcgaagaag 240
aggatgacat caccagcgcc gagtcggaga gcagcgagga ggaggagggg gagcanggcg 300
acccccagaa cgcaacccgg cagcagcag 329

<210> 90
<211> 166
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(166)
<223> n = A,T,C or G

<400> 90
tgcttcttcc ttaaagcatt tgnnacagct acagtctaaa attgcttctt taccaacgat 60
atttacagaa aagactctga ccagagatcg agaccatnnt agccaacatc gtggaacccc 120
atctctacta aaaatacaaa aatgagctgg gcttggtggc gcgcac 166

<210> 91
<211> 333
<212> DNA
<213> Homo sapiens

<400> 91
ctggctgccc accaggccgt gtatgtgagg tcaaggctga agcccggaac tgctgggcca 60
cccgtggtct ctgtgtcctg tctgtgggtg ccaacctcac cacctttgat ggggcccgtg 120
gtgccaccac ctctcctggt gtctatgagc tctcttcccg ctgcccagga ctacagaata 180
ccatcccctg gtaccgtgta gttgccgaag tccagatctg ccatggcaaa acggaggctg 240
tgggccaggt ccacatcttc ttccaggatg ggatggtgac gttgactcca aacaagggtg 300
tgtgggtgaa tggtctccga gtggatctcc cag 333

<210> 92
<211> 357
<212> DNA
<213> Homo sapiens

<400> 92
aaaagggagg tgggggtaga agtaaaagga tgatcatggg agggagctga ggggttaata 60
tatatacata catacacata tatatatttt tgtaaataaa caggaactga ttttctgcct 120
ccatcccacc catgagggct gcaggcacta caaaagagct gactactgag aattctggaa 180
aacaaggttt tttttatttg tagctatagc tacaacttgg cggcatgggg gagggtggga 240
atgtcctgga gggtctccca gccctccgca agcagagtac aaaggctgct cgggggggccg 300
gccgagggcg cgggtgcagc agtgaaagca gcagcactaa acctggtgcc cccctca 357

<210> 93
<211> 246
<212> DNA
<213> Homo sapiens

<400> 93
gctccagcct ctggggcgca ttccaacctt ccagcctgcg acctgcggag aaaaaaaatt 60
acttattttc ttgccccata cataccttga ggcgagcaaa aaaattaaat tttaaccatg 120
agggaaatcg tgcacatcca ggctggtcag tgtggcaacc agatcggtgc caagttctgg 180
gaggtgatca gtgatgaaca tggcatcgac cccaccggca cctaccacgg ggacagcgac 240
ctgcag 246

<210> 94
<211> 454
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(454)
<223> n = A,T,C or G

<400> 94
ctgaagcaag tagatgcttt ttcaaaagga aaccaaagca attgtttata tgcttggaag 60
atgtcttatt cattggaggc tgaatgctga gtctgttttt gaaaactgca tttcttgag 120
gcaggtcgca cgttctagga gtccacactg atgcaagcac agaaaagaag gaagccaagg 180

agaagtgatc ctgggggttt tctcaagccc atagttccag aaggtgcaat accagcattg 240
gtttatgatc agtctttcaa tcaacaattt gatgattagg gatctctaca ttcgtatttc 300
aggtcagaga agaacaccat ttcttgagag aagacaaaca accctagtct accaccagca 360
taggttttat catagatggg tcctgagtca gccatgattt tcttnccttc atacatcacc 420
actctaatga aacccgtctt tggcctgtgg ctga 454

<210> 95
<211> 50
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(50)
<223> n = A,T,C or G

<400> 95
tctacctttg caggaacgcg ctcatgttca acaacganct catggccgac 50

<210> 96
<211> 324
<212> DNA
<213> Homo sapiens

<400> 96
ctgtttccca aaggggtcac actcaagccc cgcagaccac acaagaatca caaaccacga 60
ggtccgtctc ccccatgact gacaccaaga cagtcaccac cccaggttct tccttcacag 120
ccagtgggca ctcgccctca gaaattgttc ctcgggacgc acccaccata agtgcagcaa 180
caacctttgc cccagctccc accggggatg gtcacacaac ccaggccccg accacagcac 240
tgcaggcagc acccagcagc catgatgcca ccctggggcc ctcaggaggc acgtcacttt 300
ccaaaacagg tgcccttact ctgg 324

<210> 97
<211> 298
<212> DNA
<213> Homo sapiens

<400> 97
aaactagtgt cagtgacact aggaatataa taaaggtaac acagcaagaa gcacagaact 60
actccctctt catctccata ttttcataat ttcttgtgtt tcaaataggg aaacatcttc 120
ctcaaagtct gcctagtgag atatggccta ctggttgcct catagctttg tacagattat 180
gaggactgaa aataattggg catttaccca tcttggtatc tgttgtatcc tttatctgtg 240
tgtgctgatt tgatcttttt tcagtttcac ataccttatc taaggtttcc caggattt 298

<210> 98
<211> 366
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(366)
<223> n = A,T,C or G

<400> 98
ctggcaggag gcccactcac tgcccaagtc atggcaacag gccggagcag cccangagat 60
gggcctaaaa tgttctggat cccttgggtc ctantgttat gttccagtct gcccacctgt 120
gctcaggatg cannnctggg atccagcacc catggaagct tctgntggga tggngtcacc 180

tatgggtttt gaaccantgt ggtatggtcc ttgggagctc tgntctgagc ttgccacact 240
gntgagagca cccactgtcc tgaccagagt ctcantggtc ctgaccccca atgngggcag 300
gggctgggca ggagggtggg gtctgctgtg ggttcagang actccacctn ctggctggtt 360
tacctg 366

<210> 99
<211> 292
<212> DNA
<213> Homo sapiens

<400> 99
cctcggaggg gcagccttcg gggcggggag cgtgagcgcg ccaaggccat ccctgagatc 60
tacctgaccc gcctgccagc agtcctcctg acatggacta tgaccctgag gcacgaattc 120
tctgtgcgct gtatgttgtt gtctctatcc tgctggagct ggctgagggg cctacctctg 180
tctcttccta actacaaaag ccctttctcc ccacaagcct ctgggttttc cctttaccag 240
tctgtcctca ctgccatcgc cactaccatc ctgtcaccag tgggacctct tt 292

<210> 100
<211> 343
<212> DNA
<213> Homo sapiens

<400> 100
tgtagtccca gttactcggg aggctgaggc aggagaatcg cttgaacccg ggaggtggag 60
attgcagtga gcccagatcg caccactgca ctccagtctg gcaacagagc aagactccat 120
ctcaaaaaga aaagaaaaga agactctgac ctgtactctt gaatacaagt ttctgatacc 180
actgcactgt ctgagaattt ccaaaacttt aatgaactaa ctgacagctt catgaaactg 240
tccaccaaga tcaagcagag aaaataatta atttcatggg actaaatgaa ctaatgagga 300
taatattttc ataatttttt atttgaaatt ttgctgattc ttt 343

<210> 101
<211> 172
<212> DNA
<213> Homo sapiens

<400> 101
aaacaatcct tgaattttcc atgttatcag aagttgttaa cagcatcgag acggaagtat 60
atgaaatata aggactgaaa taaaagtgaa tttgaaagat ggctaatcta ctagattagg 120
taaaggggga acgggtaagt ggtggggagg agtagggaac gatggggtgg tt 172

<210> 102
<211> 194
<212> DNA
<213> Homo sapiens

<400> 102
ggtcagtgat gggggagtac cactggtctg tgtgctagag gagggtgttg ctgacctgaa 60
tctgaatttc taacaggctc acagatgagg ccagcaccac tggtctgagg gccatgccca 120
ggcacacgat gttctcataa ctcggctgca atgattatat atggtggagg caagctgggg 180
ccaaggtagt tcat 194

<210> 103
<211> 342
<212> DNA
<213> Homo sapiens

<400> 103
gtgctcgggg taatgacggt gctcgaggca gtgatggtca accaggccct cctggtcctc 60

ctggaactgc cggattccct ggatcccctg gtgctaaggg tgaagttgga cctgcagggt 120
ctcctggttc aaatggtgcc cctggacaaa gaggagaacc tggacctcag ggacacgctg 180
gtgctcaagg tcctcctggc cctcctggga ttaatggtag tcctggtggt aaaggcgaaa 240
tgggtcccgc tggcattcct ggagctcctg gactgatggg agcccggggt cctccaggac 300
cagccggtgc taatggtgct cctggactgc gaggtggtgc ag 342

<210> 104
<211> 282
<212> DNA
<213> Homo sapiens

<400> 104
ctgcgtgaag atccacaacc agctcatctc gtccgtctcc aacatcacct gccccaactt 60
tgatgccagc atttgcatcc cgggctccat cacattcatg cccaatggat gctgcaagac 120
ctgcaccccc cgcaatgaga ccagggtgcc ctgctccacc gtccccgtca ccacggaggt 180
ttcgtacgcc ggctgcacca agaccgtcct catgaatcat tgctccgggt cctgcgggac 240
atttgtcatg tactcggcca aggcccaggc cctggaccac ag 282

<210> 105
<211> 297
<212> DNA
<213> Homo sapiens

<400> 105
ctggctgaga aacgagagca cgagaaagaa gtgcttcaga aggcaataga agagaacaac 60
aacttcagta aaatggcaga agagaaactg acccacaaaa tggaagctaa taaagagaac 120
cgagaggcac aaatggctgc caaactggaa cgtttgcgag agaaggataa gcacattgaa 180
gaagtgcgga agaacaaaga atccaaagac cctgctgacg agactgaagc tgactaattt 240
gttctgagaa ctgactttct ccccatcccc ttcctaaata tccaaagact gtactgg 297

<210> 106
<211> 210
<212> DNA
<213> Homo sapiens

<400> 106
ctgacagcca gcagtacctt cccaaccatt agagtgagtc accctagaag caaattctcc 60
agctccagtg catcctttag ataactgcca ctctggtcac tatcttatct acaacctcat 120
gagaaacctc agccagaacc acccagctaa gttgcctctg aattcccgag ccacagaaac 180
tgggagataa tgtttactgt ttaagacttt 210

<210> 107
<211> 338
<212> DNA
<213> Homo sapiens

<400> 107
agatggcgga cattcagact gagcgtgcct accaaaagca gccgaccatc tttcaaaaca 60
agaagagggt cctgctggga gaaactggca aggagaagct cccgcggtac tacaagaaca 120
tcggtctggg cttcaagaca cccaaggagg ctattgaggg cacctgcatt gacaagaaat 180
gccccttcac tggtaatgtg tccattcgag ggcggatcct ctctggcgtg gtgaccaaga 240
tgaagatgca gaggaccatt gtcatccgcc gagactatct gcactacatc cgcaagtaca 300
accgcttcga gaagcgccac aagaacatgt ctgtacac 338

<210> 108
<211> 426
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(426)
<223> n = A,T,C or G

<400> 108
ctgatgatgt agaagtatat gattgaacga ccagagccag aattccaaga cctaaacgaa 60
aaggcacgag cacttaaaca aattctcagt aagatcccag atgagatcaa tgacagantg 120
aggnttctgc agacaatcaa ggatatagct ngtgcaataa aagaacttnt tgatacagtg 180
aataatgtct tcaagaaata tcaatnccag aaccgnaggg cacttgaaca ccaaaagaaa 240
gaatttgnna agnactccaa aagtttcagt gatactctga aaacgtattt taaagatggc 300
aaggcaataa atgtgttcgt aagtgccaac cgactaattc atnaaaccaa cttaatactt 360
canaccttca aaactgtggn ctgaaagttg tatatgttaa agagatgtac ntctcagtgg 420
cagtat 426

<210> 109
<211> 79
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(79)
<223> n = A,T,C or G

<400> 109
aatcancaaa atttcaaata aaaaattatg aaaatattat cctcattagt tnatttantc 60
ccatgaaatt aattatttt 79

<210> 110
<211> 421
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(421)
<223> n = A,T,C or G

<400> 110
cgctggggcc tcatagttga gcacgtagta gtcgtggaca tacatgagga cggctattgg 60
ctgtccgatg atgagcgaca gccacacacc caaattggac cgcttaagag ttgcactttc 120
caaagtcaac ttctaagtct acaaggacag caacaatgtt tcagtggatt ctgaagttac 180
atgtatcaac aatttccccg gaaagctaac cctccaccgg gaactccagg tgaatgaatg 240
agtgagggaa ttcgccagat tgagttacaa agcctttcca acgattatca agagcaggtg 300
ctcggttaca acacagaggt atcctccttc acagcctttg gaccttgctg cgtggagatt 360
ttcacagata agaggggggа aatagagaga caggccttnc tccccggcca tccacacctt 420
a 421

<210> 111
<211> 274
<212> DNA
<213> Homo sapiens

<400> 111
ctgtcaacct tgacaaattg tggactttgg tcagtgaaca gacacgggtg aatgctgcta 60
aaaacaagac tggggctgct cccatcattg atgtggtgcg atcgggctac tataaagttc 120

tgggaaaggg aaagctccca aagcagcctg tcatcgtgaa ggccaaattc ttcagcagaa 180
gagctgagga gaagattaag agtgttgggg gggcctgtgt cctggtggct tgaagccaca 240
tggagggagt ttcattaaat gctaactact tttt 274

<210> 112
<211> 76
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(76)
<223> n = A,T,C or G

<400> 112
ccagagagaa gagggccagg angctgcaac aggctggcag anaggctggn cangtagtan 60
ccaccctctc cagtaa 76

<210> 113
<211> 228
<212> DNA
<213> Homo sapiens

<400> 113
cccactgaag ccgtggggac gcgcccagcg gagctaatca gattacctgg ctggtgtttg 60
cttgttctgg agtgatcttc tgactggaaa agaactatgt catggatcaa ggaaggagag 120
ctgtcacttt gggagcggtt ctgtgccaac atcataaagg caggcccaat gccgaaacac 180
attgcattca taatggacgg gaaccgtcgc tatgccaaga agtgccag 228

<210> 114
<211> 489
<212> DNA
<213> Homo sapiens

<400> 114
gtggaacaga ctgtcctcca tgtcagttcc ttctggcttc aggccctcaa ttctttccct 60
ttgagctttt ttagacccca gatctcctag gcccaggctc tctcttgacc ccagagaagc 120
cactgtcagg aaaggaagtg aaccctactg aagccagaga attcaccccg gccaaagcag 180
gccctctggg tccagccccc cattccacac cacaccagta ttgcatccat ctactgcagc 240
tacacatcct gagggcagca ccacccactc tggcctgctg gcccatcgca ggactagccc 300
aggcacctgc cgggcattgc aggatatcca gtggggcctg tgactgctcc ctgatgcgtc 360
agaagagaag tgttgcactt tagtggagga gctgaggagc acctgccccc ttgtagcttg 420
agttcctttt ggtaacagta gcagcctcca tggtggtgtc tgggacgcac gtgcacccgc 480
tgccttcag 489

<210> 115
<211> 501
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(501)
<223> n = A,T,C or G

<400> 115
ctgcaccatg ccatctatag agataggaac ggtgggtggt gggaccaacc tactacctca 60
gcaagcctgt ttgcagatgc taggtgttca aggagcatgc aaagataatc ctggagaaaa 120

tgcccggcag cttgcccgaa ttgtgtgtgg gaccgtaatg gctggggaat tgtcacttat 180
ggcagcattg gcagcaggac atcttgtcaa aagtcacatg attcacaaca ggtcgaagat 240
caatttacaa gacctccaag gagcttgcac caagaagaca gcctgaatag cccgacagtt 300
ctgaactgga acatgggcat tgggttctaa aggactaaca taaaatctgt aaattaaaaa 360
agctcaatgc attgtcttgt ggaggatgaa tagatgtgat cactgagaca gccacttggt 420
ttttggctct ttcagagagg tctcangttc tttccatgca gactcctcag atctgaacac 480
agtttagtgc tttacatgct g 501

<210> 116
<211> 452
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(452)
<223> n = A,T,C or G

<400> 116
ccatattctc atcatatcct ctctgtgtgg agtctgcctg ttgtcacaaa aaccttgacc 60
ctacatcaag ttacacctta acaaagggaa gatacaggca tcagataaaa ggtacttgtt 120
tgaaaggcag ccataaggga gaactgaact taaaaaaaaa aaaanaaaaa aaattccaag 180
ctggtttcaa cagtactttg tttccagaac aaagaaatgt ttctaaccac atcttgtacc 240
ccttcctcat caactccaga ctaccacaga ccttttttcca aaactgtgtg tcacacatcc 300
aggtcttgtg ctttanagct gcctctcagg caattttagc cagccatttc tccaagtcct 360
ggatgtcagc agagcccacg tcccctcttn cacccttggc actgcactcc angaactcca 420
ctttgagggg caactgtgan aattcaaact ct 452

<210> 117
<211> 385
<212> DNA
<213> Homo sapiens

<400> 117
aaaacattgt tttctaaaca ctaacaaaaa aaattaaggg caaactgaaa atacaaatga 60
gatttacagg cactgtgtgt agaatgtgca aaaattcact tagcttttct tttgtttttt 120
tggtgttgct ttaagaaact ttatcaaata tatttcttac aaatataaag ctttctctcc 180
caattgaagg caattaaaaa aattcaaagt ttatcaatac tcagtacaca ggtgaaccag 240
tcaaattcat tttctttctg gaaaagaata acaaaccaat atttaggatg ttcagagact 300
caacaaaaac cattctagaa atcacccaga acaattgttt tctgttgcca aagccttttg 360
ttcttcaaaa gtcaccatcc accag 385

<210> 118
<211> 286
<212> DNA
<213> Homo sapiens

<400> 118
ttggtttgcc tttttccttc ctaactttcc catatgtaga agaagccatt aagattgctt 60
actgtgaaaa gaaatgtgga aactgctctc tcacgactct caaagatgaa gacttttgta 120
aacgtgtatc tttggctact gtggataaaa cagttgaaac tccatcgcct cattaccatc 180
atgagcatca tcacaatcat ggacatcagc accttggcag cagtgagctt tcagagaatc 240
agcaaccagg agcaccaaat gctcctactc atcctgctcc tccagg 286

<210> 119
<211> 275
<212> DNA
<213> Homo sapiens

<400> 119
gtggtgaggt ttctgaagaa ttatccctga aactgccacc aaatgtggta gaagaatctg 60
cccgagcttc tgtctcagtt ttgggagaca tattaggctc tgccatgcaa aacacacaaa 120
atcttctcca gatgccctat ggctgtggag agcagaatat ggtcctcttt gctcctaaca 180
tctatgtact ggattatcta aatgaaacac agcagcttac tccagaggtc aagtccaagg 240
ccattggcta tctcaacact ggttaccaga gacag 275

<210> 120
<211> 70
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(70)
<223> n = A,T,C or G

<400> 120
cttgagactt gaaaccacaa naagtgtgan aagactggct agtgtggaag catantgaac 60
acactgatta 70

<210> 121
<211> 168
<212> DNA
<213> Homo sapiens

<400> 121
aaaagcgacc tttttgtcca ttacagaagt aacgtattta ttgtagaaat gtaatagata 60
aaaatgaaat aattattcat attctcacta ttccacaaat gtctgtgatt aacagattca 120
ttgtcaactt tagttctcat tctgcacata tgtaagttat gtttgtat 168

<210> 122
<211> 342
<212> DNA
<213> Homo sapiens

<400> 122
gtgctcgggg taatgacggt gctcgaggca gtgatggtca accaggccct cctggtcctc 60
ctggaactgc cggattccct ggatcccctg gtgctaaggg tgaagttgga cctgcagggt 120
ctcctggttc aaatggtgcc cctggacaaa gaggagaacc tggacctcag ggacacgctg 180
gtgctcaagg tcctcctggc cctcctggga ttaatggtag tcctggtggt aaaggcgaaa 240
tgggtcccgc tggcattcct ggagctcctg gactgatggg agcccggggt cctccaggac 300
cagccggtgc taatggtgct cctggactgc gaggtggtgc ag 342

<210> 123
<211> 443
<212> DNA
<213> Homo sapiens

<400> 123
aaacttactt catttattat ttgttactct ttatttctcc ctagtatgtt ttggacattt 60
gaatgtcctc ttctgtgaat ttttcatgtt tgttgcctat atctctattt tggttttaga 120
agttaaatta ttacttaaaa gaacttttta ataagtttga atgttaaatt ttgacctctc 180
atgtgcattg caaatttttt tcctcaagta tctttttctt ttttttagat agtgtttttg 240
aaagtcttca tggtgatatg cactatattc agtatatgta tgttttccta cttctcttgt 300
aaaactgttg catgatccaa cttcagcaat gaattgtgcc tagtggagaa cctctataga 360
tcttaaaaaa tgaattattc tttagcagtg tattactcac atgggtgcaa tctttagccc 420

cagggaggtc aataatgtct ttt 443

<210> 124
<211> 145
<212> DNA
<213> Homo sapiens

<400> 124
ctgaacctga gaaggaggag gcggccaagg aagaagccac caaggaggaa gaagccatca 60
aagaggaggt ggtcaaggag cccaaggatg aggcacagaa tgagggcccg gctacagagt 120
cagaggcccc gctgaaggag gatgg 145

<210> 125
<211> 391
<212> DNA
<213> Homo sapiens

<400> 125
ctgatttgtt tactgaacac tgtcacatta aatgatggtg cctaggtaaa aacgctgcac 60
acactcccct ccacccccac cccttaccca tgttgagacg tggctgcctg tcatgagatg 120
agatctgctt gagtaaagcc atatacatta cagcaagcat tccagattct taaaatgacc 180
aaacactttg gtattaatac aatgtattcc ctgttttctc aaatatacaa aatatacatt 240
tccagtttta gttgtggttt tcttgttttt ttttgttttt gttgttttta cacaggaata 300
gttaggtctg tcatttgagg gagcccaggg gacctggaac gggtcacacg ggcagtgctc 360
agttctggtg cctcttcata tgcagggcca g 391

<210> 126
<211> 306
<212> DNA
<213> Homo sapiens

<400> 126
aaaaatcact acatcaaatg ggatagagag taagaagaca ggagagagag gagaaaccat 60
gttttttgtt ttgagtcagg agggtctcac tctgtcactc aagctgaagc acagtggcac 120
aatcacagct cactgcagcc tcaacctccc aggctcaagc gatcctacta cttaagcctc 180
ccaagttgct gagactacag gcacaagcca ccatgcccag cccaatttga ttgtgtttca 240
tacagatagc cagttttccc agcaccaacc cggacttgtt aaatagcctg ttcttttctc 300
actttt 306

<210> 127
<211> 153
<212> DNA
<213> Homo sapiens

<400> 127
aaaaaatccc acttttcgaa aatatctgac aatcaagggc acagagacta gcgtaatgct 60
gattctcact ggcgcaaaca gcttgtggat cgcataggcc accacgaagg tacttgtgcc 120
tgctgccatt tttgactgta ccagggactc ttt 153

<210> 128
<211> 134
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(134)
<223> n = A,T,C or G

<400> 128
gctttcattc ctgttcanaa ntcaatgccc ttgacggggc tgatgtgtnn agctgntaac 60
anncacccat cccagtgtca ggangatttg annnaggagt ttggangaga gtgggaagga 120
atgactgctt anga 134

<210> 129
<211> 246
<212> DNA
<213> Homo sapiens

<400> 129
aaaggctttc attcctgttc agaagtcaat gcccttgacg gggctgatgt gttgagctgc 60
taacagtcac ccatcccagt gtcaggaaga tttgatagag gagtttggag gagagtggga 120
aggaatgact gcttaggagg ggagagagcc tggcaatgaa atgtggccca gggcaccagc 180
ctgacagccc cgagggaccc ctgggtgtgt ttgaggcttt catagttcag atttctgcat 240
gcccgg 246

<210> 130
<211> 460
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(460)
<223> n = A,T,C or G

<400> 130
cacaattcta ccgttcattt ntgtaactgc tttagtggct tcttctgggg aggagaaaca 60
tacaaaacca aacccttgc tgcgaccacc ctccatcata acctttgcac tagtgattgn 120
accaaatgga gaaaactctt tccggagacg ttcatcatca ataccatcat caagattttt 180
cacataaaga ttaacaccct ggtatctggt gatcctatct tgtttcatct gttcaaattt 240
gcgcttannt tccgtctgcc gttccacctt tttctgagct cgaccaacat aaatttgttt 300
tccattgagc tcctttccgt tcatctcatn cacagctttn tgtgcatctt catgcctttc 360
aaagcttaca aatccaaatc ctttggattt tncactttca tcagtcatta ctttcacact 420
taaggcaggc ccaaacttgc caaagagatc cttaaggcgc 460

<210> 131
<211> 464
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(464)
<223> n = A,T,C or G

<400> 131
tgacctgnat ctctctgcta ttaatgacaa aagcatcgnc aaaaagacnc cacagttagc 60
aaaaacaata tcaaagaaac ctgagtcaac atcattttct gcccctcgga aaaagagccc 120
ggatttatct gaagcaatgg aaatgatgga gtctcanaca ctactgctga cgctactatc 180
cgtaaagatg gagaacaatc ttgctgagtt tgaaagaagg gcagaaaaga atttattaat 240
aatgtgtaag gagaaggaga agctacagaa aaaggcccac gagctgaagc gcaggcttct 300
cctctctcag aggaagcggg agctggcaga tgtcctggat gcccagatcg agatgctcag 360
ccccttcgag tgcgaggcga ggtggctcct gacgctaana ggtgtgcatc accttcgacc 420
aggccaacct gaccgtcaag ctgccagatg gatacnaatt caag 464

<210> 132
<211> 303
<212> DNA
<213> Homo sapiens

<400> 132
ctgcggtggt caggtcccgg tattcccggt acatgttgtg ggtgccgctc cgggagtcat 60
agcgcagcca gatcccgaag ttcttcaccc gcaggggga cttctcaaac acctgcccac 120
agtagacaat ctcccctgaa gacttcttca tcttctttaa ctgagataca aagtaccaga 180
agcgggactt ggcgacgaca tgattaggcg caaagattcg catgcggtag aggggcggcg 240
tgtggcattt gggggtgggc aggcagcgac ccactacctt gtactctcgt agcgtgcccg 300
agg 303

<210> 133
<211> 273
<212> DNA
<213> Homo sapiens

<400> 133
gtggatgatg tctgtggcga tggcattcaa gaagttgcgg tcgtgggaga cgactaggat 60
ggtggagggc cacgtctgca ggtaattctc cagccacagg atggccctga catccagcat 120
gtttgtaggt tcatctaaca gcagaagatc tggcctagca aagagggccc gggccagggc 180
cagcctcatc ctccagccac aaacagcacc attgtattgt tgaatgttta tgtaactgat 240
ggcttttcta taatgtaatt tttgaatgtt cag 273

<210> 134
<211> 507
<212> DNA
<213> Homo sapiens

<400> 134
ctggtccttc aggcaaaatt tggaggtcac aatgaactcc aagcctgaca caaagatatt 60
ctacagtttc acagctatca tttgtacata ttaagttgat tcactctttt tgagcaaatc 120
tacctagaaa acggcaaatt aatatattcc tttacataca actttgtgtc tcaaaattct 180
tgaaaaacaa gagcagatga ctttgtattc aaagactacc aaagtatgta tttgattttc 240
acatgcaaac aacttaaaac cttataaatc tcatgtcaac tctgcatgat gccttgaagg 300
aaatgacata caaagtttgc taactgtgca aaatattaaa ttgctaaaac attttacata 360
atgaaataat acatgtaaat gttgaagttg acacatgaaa ttaacatggc ataagaactt 420
atcacatttc agatattttc tttagtaaca agtttttgtt tttatagttc ctggtacaca 480
gcaaagttta tcacgaaaga taaaaat 507

<210> 135
<211> 148
<212> DNA
<213> Homo sapiens

<400> 135
ctcggcggcc acagacatca cgtcctccct atcagacgac caggtacccg aggctttcct 60
ggtcatgctg ctgatccagt tcagtaccat ggtggttgac cgcgccctct acctgcgcaa 120
gaccgtgctg ggcaagctgg ccttccag 148

<210> 136
<211> 150
<212> DNA
<213> Homo sapiens

<400> 136
ctgctaagaa gcagacattg tctctacaag ctcagagaga agagaaagca aaagcctccg 60

agctctccaa aaagaaagca tctgccctgt tgttcagcag tgatgaggag gaccagtgga 120
atattcctgc ttcacagacc cacttagcat 150

<210> 137
<211> 179
<212> DNA
<213> Homo sapiens

<400> 137
aaatgcactg ttctggttcc taacttgaag cagttgtcct tgtgagaacc ggtctttgcc 60
tttagctcat gtcgtgtttc acagcaaaga gggtacagaa ccatcactgg tccaggttaa 120
tgtacaaaat tttctggcaa tgcctgatta aaaaaataaa attggcttgt tgagaacag 179

<210> 138
<211> 249
<212> DNA
<213> Homo sapiens

<400> 138
ctgcactgga agcttccagg gatgttgatg cagcggtagg agcagatgtg gcccccggtg 60
ggcagggcgc actcgtcgat gtcttcacag gtgactccat ccacatcgct gagctggtag 120
cctcgccggc agtaacactg gtaggagccg tagacgttgg cacactcctg gctacagggg 180
ctgctgctgc actcattgat gtcttcacat gacctgccat ccacagagag ccggaagccc 240
acggaacag 249

<210> 139
<211> 237
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(237)
<223> n = A,T,C or G

<400> 139
aaaaccatca taacaaaaag ggtccattgt cttatgatcc actggaaaga ggaccgactc 60
atcatttatg gctatgactt ggcagtgact ccaatgtgat atcctgtaat tttatcttca 120
gttatgctat agcatgtaca tttccattct cttgtcgaag tttctttcgt tcctcanctt 180
ctccttcata tttcctgacg tattgtcttc taagctggac tgtaataaca gcaacag 237

<210> 140
<211> 342
<212> DNA
<213> Homo sapiens

<400> 140
cttccatcat gaaacgggat gacagcaaca ataagacttt ggctgagcaa aacactaaga 60
atcctaaaag cactactggt agaagttcca aatctaaaga ggagccatta tttccattta 120
atttggatga atttgttact gtggatgagg ttatagaaga agtgaatcct tctcaggcca 180
agcagaatcc actaaaggga aaaaggaaag aaactctcaa aaatgttcct ttctctgaac 240
ttaacttaaa gaagaaaaag gggaaaactt ccactcctcg tggtgttgag ggagaactat 300
cttttgtgac attggatgag attggggaag aggaagatgc ag 342

<210> 141
<211> 226
<212> DNA
<213> Homo sapiens

<400> 141
gtcctctaga gaatcccctg agagctccgt tcctcaccat ggactggacc tggaggatcc 60
tcttcttggt gtcagcagcc acaggagccc actccctggt gcaagtggtg cagtctgggg 120
ctgaagtgaa gatgcctggg gcctcagtgc aggtctcctg caggacttcc ggatacacct 180
tcaccgacta ctacatacat tgggtgcgac aggcccctgg acaagg 226

<210> 142
<211> 235
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(235)
<223> n = A,T,C or G

<400> 142
ccagcgacct cccggttcaa ttcttcagtc cggctggtga accaggcttc agcatccttc 60
cggttctgct cggccatgac ctcatattgg cttcgcatgt cactcaggat cttggcgaga 120
tcggtgcccg gagcggaatc cacctccaca ctgacctggc ctcccacttg gcccctcagc 180
gtactgattt cctcctcatg gttcttcttc aggtaggcca gctcttnctt cagga 235

<210> 143
<211> 508
<212> DNA
<213> Homo sapiens

<400> 143
ctgacaaaaa gtgtcagagc cagaggccaa cctctgctag atgaagcagc agcacatgac 60
tccatttcta tctgataagg agacagagaa gaggcatctc gaacagatga aaaaccaaag 120
gctggtgtcc taaaaaaaca gattggcttc aaagaaaaca ctaaggaaga cccacagagc 180
tgtattaatt ttagtaaaaa taatcatatg ccaacagggg aattgaacca ctttctaaat 240
catagtatga actcatctct tcagatactt ggtaagtggt caaagcttgt ttttataatt 300
actttcactg tcttgggcaa aaagtctttc ttatctttgg tccttaggtg tggtatcagt 360
ttcttccatt tttttatgtg ttacaaaaca atctttttt tacttgacat caacaaccaa 420
ggtgcagtat aaacacgaag ttgctgatat tgttgctttt atacacataa aataccaaca 480
tctcccatac attttatagg ctatacga 508

<210> 144
<211> 382
<212> DNA
<213> Homo sapiens

<400> 144
cctgccgtcg atgccaggga ggccgacagg accttctttt ccagcggggc cgatatttcc 60
aggggaacca ggaagacctc tgggtcccat gagaccaggc tccccagggc gaccagcatc 120
tccattaggt cctcggactc cagcagggcc acttgcacca cgactaccag gagggcccat 180
gacgccagct ctgccatcag ctccaggaag accacgagaa ccaggactac ctctcagccc 240
aggaggtcct ggagggccgg cagatccagc ttccccatta gggcctctct ttccttcttc 300
accactggga ccaggaggac cttggggccc agcagagccg ggctcaccct tgttaccgct 360
ctctcctttg gagccagacc tg 382

<210> 145
<211> 109
<212> DNA
<213> Homo sapiens

<400> 145
gctaacatgc cttggttcaa gggatggaaa gtcacccgta aggatggcaa tgccagtgga 60
accacgctgc ttgaggctct ggactgcatc ctaccaccaa ctcgtccaa 109

<210> 146
<211> 87
<212> DNA
<213> Homo sapiens

<400> 146
gtgaagtacc acggagaaat catattggaa agttactact tagccatctg acttgacttc 60
cttggttatc aaataattac atattct 87

<210> 147
<211> 396
<212> DNA
<213> Homo sapiens

<400> 147
aaagataaaa ccaacatgtc cagtgctatc cttatgcatg gtaatcgtcc gttcaaaggg 60
cctgtcacga atggtcatgg taatcttctc tccaaaagcc tgtttgagca ccttgtgcgc 120
tttatcagag ctccatcctg cacagttttc accattgatc tgaagtactt ggtccccaaa 180
tctcagacca accaatgagg ctggagaatt agcctggact agctgaacaa atataccatt 240
atctattgat ttaagcctga gtccaatttc tccatcttga tccttacaca aaatgacttc 300
acgaatccct tgcttaattt ctgctctacg aattccaaca tcattaccag ttacaggagc 360
caccatatag tttatactgg aaggtcttgc taccaa 396

<210> 148
<211> 503
<212> DNA
<213> Homo sapiens

<400> 148
aaatcccaat ttcccatctt catcttcaga aaccatttca aacgtatcaa actgtaattt 60
cttcataaca gccacatatt tttcttcaag tgactttaat actgacaaag gtttgggttt 120
catagccgcc ttcttggagt attcacccag tttttttttcc tgatttgctt gccgcaaact 180
ggtggtggct gcataaacta tctcagcagt cttttggatg tctggtacca aaagagtaag 240
tccttctggt tcttgatcag acgcatctgg ttttactcct gttttaacat tttccctttt 300
agatcttaaa cggttggtat aggtatcaac acaggtcttc attttggcta acaatgtacc 360
aacagaagtt tgacattctg actgttcttc ttcctcttca ccgttctctg tagaaagggg 420
caacaatagg ggcaccatgg cagcacaaga agcaatggcc cgaagcaatt ccagcagtgc 480
ccgatagagt ggcacatgtc ttg 503

<210> 149
<211> 196
<212> DNA
<213> Homo sapiens

<400> 149
ccattaaaag ttatttacaa cagtgggaga aaaaaagaca agaagttgtt tcacactaca 60
gacctccccc caccccaaag cctaatactt gcttaccaag tcaaaaaaga gacacagttg 120
attcacaggc tggaggtttg aacttgagta agacatttat aaaaacctag acggggcagt 180
gtcctcccca gcccag 196

<210> 150
<211> 147
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(147)
<223> n = A,T,C or G

<400> 150
ccttctctga aaaaagagaa ggaattactt attaaaacta agcacactta gcaacttctt 60
tccnatccta tctttattcg tttgcctngt gccaaatttt tctngccctt tttaatttgc 120
aaaccttnaa aaaaaaaacc aaaaaac 147

<210> 151
<211> 419
<212> DNA
<213> Homo sapiens

<400> 151
ctgcgctatg gcgaagacgg cctggcaggc gagagcgttg agttccagaa cctggctacg 60
cttaagcctt ccaacaaggc ttttgagaag aagttccgct ttgattatac caatgagagg 120
gccctgcggc gcactctgca ggaggacctg gtgaaggacg tgctgagcaa cgcacacatc 180
cagaacgagt tggagcggga atttgagcgg atgcgggagg accggcaagt gcgtgtcctg 240
ctcttcagaa gtggagtgaa gggcgtgttc tgtgcaggtg cagacctgaa ggagcgggaa 300
cagatgagtg aagcagaggt gggggtgttt gtccagcgac tccggggcct gatgaatgac 360
atcgcagcct tccctgcacc caccattgcg gctatggatg ggtttgcctt gggcggagg 419

<210> 152
<211> 241
<212> DNA
<213> Homo sapiens

<400> 152
gtgccagtca agatgcctgg ctcaggccat caggagctgg ttagccccat tccaccccca 60
gccctgcatg cagggtccag ccattgtctt tgggggaaac aggcagaata agtggaggat 120
ggagctgggg cttgggctcc tctaggtacc ttctgagagc tttgacaagc cagaaagaag 180
ctaccaggtt gagggtgctg gtcttctgga ctcaggagag acatgttcgc cgaggatatc 240
a 241

<210> 153
<211> 271
<212> DNA
<213> Homo sapiens

<400> 153
ctgtctcacc agctccctaa ctcatgtgta cctgcacctt cctcttgaaa tctgaacatt 60
ataataccac aagccacttt cagcctccag tgggaaggct ccagccacac gccgatattt 120
cgtcctgctt cccgtcatct catatctaaa agtcatggct taagttaggc aataaaacct 180
gtggctttag gcatctttag taaaaaagct gaacaaatcc caaatttatt cccattttct 240
tgagaaataa acttcataaa acaacagaca g 271

<210> 154
<211> 120
<212> DNA
<213> Homo sapiens

<400> 154
ccatggcgct cgggtgcgcg cagtgcacgc gggttatcac cggagtggga ctggtgactt 60
cattagaaga ggaaggaaga cctgagctgg cctgtgaata tgctccgccc cctgcatcag 120

<210> 155
<211> 92
<212> DNA
<213> Homo sapiens

<400> 155
ccatggccca ggtcacccac cccctggtcc acatcactga ggaagtagaa gagaacagga 60
cacaagatgg caagcctgag agaattgccc ag 92

<210> 156
<211> 501
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(501)
<223> n = A,T,C or G

<400> 156
gtgtgagcca ctgcaccagg caaactgcga tcttttagng gtgcctnttc tctcttttga 60
cttaaggatg ttgtccctta aggaaacctg gaggctacta ctgtgataca ctacttgaga 120
gatggattgt tgcgctttct tctacagtct ttacaaggag tagattataa agacagaaga 180
tgttaaccat tgcattaatg tttggaagct gacagtcttc tagatttctg ctagcaaact 240
gatatgaggt agagtcctga aagatctttc agcaatttca ttttcttggg ataagtgagt 300
cactttcaga acagtatgtg ttgtagaatt ttttggttgt ggctgctcta ctcagattgc 360
atagaggttt ttttgntttc tgntttctgn ttgnttgntt tggtcagatt ttttgaaaca 420
tcctcaaagt gactattcag ttttcaggat gatacactat gaagatgttt caaaaaatct 480
tcatagtgta tcatccacct c 501

<210> 157
<211> 527
<212> DNA
<213> Homo sapiens

<400> 157
aaaggagcca gcaccatagc agagtacata agtggctatc agagaagcca gccgatatgg 60
attggcctgc acgacccaca gaagaggcag cagtggcagt ggattgatgg ggccatgtat 120
ctgtacagat cctggtctgg caagtccatg ggtgggaaca agcactgtgc tgagatgagc 180
tccaataaca actttttaac ttggagcagc aacgaatgca acaagcgcca acacttcctg 240
tgcaagtacc gaccatagag caagaatcaa gattctgcta actcctgcac agccccgtcc 300
tcttcctttc tgctagcctg gctaaatctg ctcattattt cagaggggaa acctagcaaa 360
ctaagagtga taagggccct actacactgg ctttttagg cttagagaca gaaactttag 420
cattggccca gtagtggctt ctagctctaa atgtttgccc cgccatccct ttccacagta 480
tccttcttcc ctcctcccct gtctctggct gtctcgagca gtctaga 527

<210> 158
<211> 323
<212> DNA
<213> Homo sapiens

<400> 158
ccacttacac ttgtgaccag tgtggggcag agacctacca gccgatccag tctcccactt 60
tcatgcctct gatcatgtgc ccaagccagg agcgccaaac caaccgctca ggagggcggc 120
tgtatctgca gacacggggc tccagattca tcaaattcca ggagatgaag atgcaagaac 180
atagtgatca ggtgcctgtg ggaaatatcc ctcgtagtat cacggtgctg gtagaaggag 240
agaacataag gattgcccag cctggagacc acgtcagcgt cactggtatt ttcttgccaa 300
tcctgcgcac tgggttccga cag 323

<210> 159
<211> 541
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(541)
<223> n = A,T,C or G

<400> 159
ctgctatgtg gtggccgctg tggctgacac tgagtgaagg tgtttgaaat gcaggagagg 60
atatcccagc aaattgggat cacatgcttt tgtctccaca gcaaccagcc actgcaggca 120
gcatgtcttt cctcccctgc tctctgcttg ctgttgtttt gacgctattc tgcttgcatg 180
tcttctggtt gggatgtgga gttgttgctg gactctcagg cgaagctgaa gtcattgaag 240
tgtgtgaagc tctgtgcttg catgagggca agcaaggaat ggctgtgcct gaggctgctc 300
tgggaaactc cttgcccctt gacctctttt gagagcattc acgtggtctt cttgctcatc 360
cccttataaa tgtgctttgc ctgcctcagc ctcatggtca gagcagtgga gactggagcc 420
ctgtttgcac gttctagttg ttcggagaaa gcctaggttc tgggctcang tccagatgca 480
gcggggattc tgttctctga ctgtggcgac cttgctttgg ttcttgttga agtgaaccaa 540
g 541

<210> 160
<211> 378
<212> DNA
<213> Homo sapiens

<400> 160
cctgggagat cccagggtcc tccaccctcc ccctgaccac atacaaaggc actctagttc 60
aagggtgaaa agtctcaccc aggaggaaca gccctccttg aagcaatggc agggccagca 120
gggaggtggg catggcaggg aatggagtga gccagacaga cttcacctcc ttactggaca 180
cagggtcaag ggcgagtttc aattgctgct ccctttactt tctctacctg tgactactcc 240
ctggaccaat cctgaggagg gcacattttc cagaagccac gtgatagggg ctggtttctg 300
tggagccgga ggcagagaca ctgaacttga gctcacctcc taacaccggc agtaaacttc 360
ctggaacttt gccctcag 378

<210> 161
<211> 388
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(388)
<223> n = A,T,C or G

<400> 161
ctgaagaaga agctgccgac ctcccaacaa agcctacaaa gatctccaag tttggatttg 60
ccataggtag tcagacgaca aagaaagcat cagccatatc catcaaactt ggatcaagta 120
agcctaaaga aactgttcca aatattgaac aacagggtgg acagatgatc atgaagcaac 180
tatatatgtt agatgacaag aaagaacctc ataattcaca gtcattttcc aataaatgtt 240
tatgatgagt tttgatttct catgatttcc tttataaatt ccccaggata aactaagttg 300
ctctangatg agcttgggaa gctaggttaa aacaggaacg aggcatcaca ggatagaaac 360
aatcctggtg ggattcacct atcaccag 388

<210> 162
<211> 300

<212> DNA
<213> Homo sapiens

<400> 162
ctgccaaaat ctgctggaat cctttgatgg tctccttcag gggtaccagc ttccccatat 60
gacctgtgaa gacctcagca acctggaatg gctgagacaa gaaacgctgt attttccgtg 120
cacgggacac ggtcaacttg tcttcctcag aaagttcatc catacccagg atggcaatga 180
tatcctggag ggatttgtag tcctgcagga tcttttgcac cccacgggca acatcgtaat 240
gctcactgcc aacaatgttg ggatccatga tacgagaggt ggagtctaga ggatccacag 300

<210> 163
<211> 197
<212> DNA
<213> Homo sapiens

<400> 163
aaaactacaa acacaatatt gactaaaaaa ggaaaaaaaa gggaataaca tgtatctaat 60
aaaatattca gtataaaaag aggactaatg gaattaagtg gccccttcc ccatttttac 120
attctaaaca atgattccat caagacaaat cattaaaaag tgttattaca ctgatttttt 180
ttttttaata agaagga 197

<210> 164
<211> 548
<212> DNA
<213> Homo sapiens

<400> 164
cttcttttgg tggtaaatag gtatttattt gaaatgaaaa aaaaattact taagtacctg 60
gactattgca tttaatcatg tattgtaatt gtgttactct acctttttgc atcagagaca 120
aatatacaat gaacattcag atatcacaga ctgcacacta gatagtaatt cttcaggtct 180
tttacataac caccaagaaa cagatattgg tttctgcaat atagtataaa agtccacaat 240
caatccagtc ttagccagta tcttcaattt acttctgttg ctgtacaaat aattggagaa 300
gggctttcct tgcagaggaa atacatggac tgtagaagat actcctcagg gtgtcaggag 360
gtgaaaatga agcttctgag gtttgcaaga aaatgtttac aaataagagt ctggcattta 420
gtatcctcgc atgcatctcc agcatgggaa actataacac ggctggcccc aggctcgtcc 480
tgtctggctg cctctttgta agaggggaga agattgcaca gtgtgatgga gctcattttc 540
agcagagt 548

<210> 165
<211> 485
<212> DNA
<213> Homo sapiens

<400> 165
aaacagaacg agacaccagc taggattata actttagcat tctatagcag tctgctcaca 60
cagccctcct ccatgctggc tcttgggcca cactgttccc acatggagct tgagtctcct 120
ccaacacatt ccatgagctt caagtgcaga gacatggtgt acacttcggg ctgttctaca 180
gagcactcca gaccatacgt ggctgaatac gtgagtgagt gtttttctgt ccacttataa 240
accatgttga tattaagcat aaatataatc caaatcagct ttccttttct tggcctaaag 300
gaatatgatg ggattaaaac agaagtgaat taagcaaaga tccactattc tgaacaaata 360
acatagaagt gattgaacaa tttggaccca ctaaattttg tgtctagctg taaaatggac 420
attgtgataa aaacaggatt tgaaggaaaa tgaatagcta atttgtcaat taaataatta 480
aaact 485

<210> 166
<211> 198
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(198)
<223> n = A,T,C or G

<400> 166
agcgtggtcg cgggtcgagg tntgccaccc ggctcttctt aacctgtttt gttttctgct 60
cagcacggtt aaaagaccaa cgtgtgtgga tcaaatataa aggccacacc tttcagaccg 120
aacctactca aagatccttt actttgcaat agtttgaact ggagaaccaa agacgggaga 180
cgaatgaaag caaagatg 198

<210> 167
<211> 539
<212> DNA
<213> Homo sapiens

<400> 167
ctgtttcata ggctggagat gcactcttct agactgctcg agacagccag agacagggga 60
ggagggaaga aggatactgt ggaaagggat ggcggggcaa acatttagag ctagaagcca 120
ctactgggcc aatgctaaag tttctgtctc taagcctaaa aaagccagtg tagtagggcc 180
cttatcactc ttagtttgct aggtttcccc tctgaaataa tgagcagatt tagccaggct 240
agcagaaagg aagaggacgg ggctgtgcag gagttagcag aatcttgatt cttgctctat 300
ggtcggtact tgcacaggaa gtgttggcgc ttgttgcatt cgttgctgct ccaagttaaa 360
aagttgttat tggagctcat ctcagcacag tgcttgttcc cacccatgga cttgccagac 420
caggatctgt acagatacat ggccccatca atccactgcc actgctgcct cttctgtggg 480
tcgtgcaggc caatccatat cggctggctt ctctgatagc cacttatgta ctctgctat 539

<210> 168
<211> 555
<212> DNA
<213> Homo sapiens

<400> 168
ccatctgatc tataaatgcg gtggcatcga caaaagaacc attgaaaaat ttgagaagga 60
ggctgctgag atgggaaagg gctccttcaa gtatgcctgg gtcttggata aactgaaagc 120
tgagcgtgaa cgtggtatca ccattgatat ctccttgtgg aaatttgaga ccagcaagta 180
ctatgtgact atcattgatg ccccaggaca cagagacttt atcaaaaaca tgattacagg 240
gacatctcag gctgactgtg ctgtcctgat tgttgctgct ggtgttggtg aatttgaagc 300
tggtatctcc aagaatgggc agacccgaga gcatgccctt ctggcttaca cactgggtgt 360
gaaacaacta attgtcggtg ttaacaaaat ggattccact gagccaccct acagccagaa 420
gagatatgag gaaattgtta aggaagtcag cacttacatt aagaaaattg gctacaaccc 480
cgacacagta gcatttgtgc caatttctgg ttggaatggt gacaacatgc tggagccaag 540
tgctaacatg ccttg 555

<210> 169
<211> 193
<212> DNA
<213> Homo sapiens

<400> 169
ctgcggccca tgatgtcaga gctggaagag agggcacgtc agcagagggg ccacctccat 60
ttgctggaga caagcataga tgggattctg gctgatgtga agaacttgga gaacattagg 120
gacaacctgc ccccaggctg ctacaatacc caggctcttg agcaacagtg aagctgccat 180
aaatatttct caa 193

<210> 170
<211> 207

<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(207)
<223> n = A,T,C or G

<400> 170
aaaggcagac actgagtcag tattaataga ttaactaaac tgcactgtaa tttagataaa 60
attactgtgt ctcactgtnt attacatgca aaatccacat aaattgtcat ttaaccaaca 120
gtactgnacg agcgaacatc tcgatatatg aaaactgcat catcaattca acgttttggt 180
acttgaaact gcatcataaa tgcaaca 207

<210> 171
<211> 265
<212> DNA
<213> Homo sapiens

<400> 171
cctggcttcc ctgccagtcc ctgtccttca cactatgagg gagagtcctg acttgaaatc 60
agaagacctg agcatctatt cttggctctg ccacttatta ttgtgtgacc aataatctct 120
ctaggtttca gttacctcct tcataagtgc tctgtgcagt aaggaaggag aggggaagca 180
atggtctgtg gtgctaaggg agagccagat ggtgctggtg tctgaaggag gagggagaat 240
attctgagca ggggcaatga tgtgg 265

<210> 172
<211> 449
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(449)
<223> n = A,T,C or G

<400> 172
ccatgattct gtcttttcaa tgactgtggc ttctactcna acaanatcct tnncnaggag 60
tggcttgcca agcagnntga agttgtctgc cccaaccagc aggaccttnn ccagtcgaat 120
tnnctctcca cacgcaaggn ctanttcatt tccaattaan atcaggtctt cagaggtcac 180
cttccactgg cggctggcaa agtgcaccac ggcaaatagc ctgccatact gccccgtgac 240
gatcatctca ttcaccttct tcacgacctc tgcatggtgt ctggtctcct caactgggtc 300
tggcagaaca acttctggac anngtggtga actcagggat gttttaggaa catatcctgg 360
tanatatgaa gtgctctgtg aattgaacct tcgagaanca gaccaaaggg aggctgctcc 420
gggccccgaa ggtctcanga tgctgtggc 449

<210> 173
<211> 367
<212> DNA
<213> Homo sapiens

<400> 173
cgagcggccg cccggcaggt ccattggcgt aaaccttgaa gcgatccaag ccacagcgaa 60
tggacagatc aaagaactgt ccgggaccaa atgggttgtg ggtgatcttc ttctcctcgg 120
atccccacga gccattcaga aggctgttcc ggaccacggt accgttgccc atgcggggat 180
taatgtgcag agctatgtcc cctgaggagc ccaccttgaa gttgatagca aagctcttgc 240
ctgtgggagg cacatagccc ttgatgatga tggttcttcg agctgtgagc cctccttgca 300
gcctcccgaa atatggcaca ggcgggttga aggttggggg tccttccatg gtgggcaggc 360

tgttcag 367

<210> 174
<211> 458
<212> DNA
<213> Homo sapiens

<400> 174
ggcagccatc tccttctcgg catcatggcc gccctcagac cccttgtgaa gcccaagatc 60
gtcaaaaaga gaaccaagaa gttcatccgg caccagtcag atcgatatgt caaaattaag 120
cgtaactggc ggaaacccag aggcattgac aacagggttc gtagaagatt caagggccag 180
atcttgatgc ccaacattgg ttatggaagc aacaaaaaaa caaagcacat gctgcccagt 240
ggcttccgga agttcctggt ccacaacgtc aaggagctgg aagtgctgct gatgtgcaac 300
aaatcttact gtgccgagat cgctcacaat gtttcctcca agaaccgcaa agccatcgtg 360
gaaagagctg cccaactggc catcagagtc accaacccca atgccaggct gcgcagtgaa 420
gaaaatgagt aggcagctca tgtgcacgtt ttctgttt 458

<210> 175
<211> 325
<212> DNA
<213> Homo sapiens

<400> 175
ccttctcatt tgaggggatt cctcaagact caaccccaca ggcccccact gtaggaaaca 60
agccagagaa agcagcattc agagaatggg ggacagagaa ggggaaagat atgatcccaa 120
atgcagtaca aagttggcgt ctggttctga cacaaaccag atactgaagc actcacggtc 180
aggtcagcaa cctcctttga tggaccccca aaagctgact gaccaggcaa actgctttca 240
aggaatgaaa gagtggaggg tagggcttgt agcaaacaag ccagtttcag tcactctgtt 300
cccccaggag aacaaccttt agcac 325

<210> 176
<211> 195
<212> DNA
<213> Homo sapiens

<400> 176
gtggtctgag ctcggcctat gggggcctca caagccccgg cctcagctac agcctgggct 60
ccagctttgg ctctggcgcg ggctccagct ccttcagccg caccagctcc tccagggccg 120
tggttgtgaa gaagatcgag acacgtgatg ggaagctggt gtctgagtcc tctgacgtcc 180
tgcccaagtg aacag 195

<210> 177
<211> 214
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(214)
<223> n = A,T,C or G

<400> 177
ctgccacccg gctcttctta acctgttttg ttttctgctc ancacggtta aaagaccaac 60
gtgtgtggat caaatataaa ggccacacct ttcagaccga acctactcaa agatccttta 120
ctttgcaata nttngaactg gagaaccaaa gacgggagac gannnaaagc aaagnngctc 180
aaagaaccaa aggaaagacc tgaaggaatc caca 214

<210> 178

<211> 310
<212> DNA
<213> Homo sapiens

<400> 178
cctgtgggct tttcccaaca agcaggctca gtgccagcct ctgtgtcagc ctccagggca 60
cgccaacctt ctcatggtgc cccaagcccc accccaatgc acacatagga agtctccagg 120
ctgcttgggc agaggcacaa tcattttaga ttaaaaaaaa ttgaacaaag agaccctctt 180
gcgagaggtg agatgaggcc ctgccatgca aaggagtccc agcagaggag gaagaattcc 240
atcctggagt tcaagtttct gtgcagagac aggacctggg gacagagaac ggtcctccac 300
ccaatttcag 310

<210> 179
<211> 386
<212> DNA
<213> Homo sapiens

<400> 179
cccgccttcc ccggtcccag cccctcccag ttcccccagg gacggccact tcctggtccc 60
cgacgcaacc atggctgaag aacaaccgca ggtcgaattg ttcgtgaagg ctggcagtga 120
tggggccaag attgggaact gcccattctc ccagagactg ttcatggtac tgtggctcaa 180
gggagtcacc ttcaatgtta ccaccgttga caccaaaagg cggaccgaga cagtgcagaa 240
gctgtgccca ggggggcagc tcccattcct gctgtatggc actgaagtgc acacagacac 300
caacaagatt gaggaatttc tggaggcagt gctgtgccct cccaggtacc ccaagctggc 360
agctctgaac cctgagtcca acacag 386

<210> 180
<211> 304
<212> DNA
<213> Homo sapiens

<400> 180
gtggagttac tggcctactc cttccccatg agccctccct gtctgcactg cccaggccag 60
agggtagagc acaggggttt ccccatacta cctccccact gggtccagtc ttgacaaagg 120
caggaagcca gctagggtgg gggcgatagg gtcagcgggt atgtcccact gttggaggtc 180
actggtattc tgtttgtttt tgttttgttt cgttttgttt tttgagacag ggtctcgttc 240
tgtcgcttag ctggagtgcg gtggcgtgat catggcactg ctattcttga agcactccac 300
ccac 304

<210> 181
<211> 341
<212> DNA
<213> Homo sapiens

<400> 181
ctgcctccct tgaaactctc ttcccaatca agggctccca aggagctgca ggccaagtcc 60
tctgctccta tttagcaaga ggcaggcggc aattcgggct gatctcccca tcacccttca 120
tttaaccgca aaaaagtcac caaccaactt ctcagacccc ctgggcaatc cagggtttct 180
tgtttcctaa gctcctatgg aacaagcaat cagttctttc ttggactttt ggttcaattc 240
cttctcattc agaggaaata tggttgccgt gtaggcagat gtctcctagg agcgtgtgtg 300
tgtaagagcc tgtgtgaaat tcagccaggt tagcaccaag g 341

<210> 182
<211> 533
<212> DNA
<213> Homo sapiens

<220>

<221> misc_feature
<222> (1)...(533)
<223> n = A,T,C or G

<400> 182
ctgaacaaca atggctatga aggcattgtc gttgcaatcg accncaatgt gccagaagat 60
gaaacactca ttcaacaaat aaaggacatg gtgacccagg catctctgta tctgtttgaa 120
gctacaggaa agcgacttta tttcaaaaat gttgccattt tgattcctga aacatggaag 180
acaaaggctg actatgtgag accaaaactt gagacctaca aaaatgctga tgttctggtt 240
gctgagtcta ctcctccagg taatgatgaa ccctacactg agcanatggg caactgtgga 300
gagaagggtg aaaggatcca cctcactcct gatttcattg caggaaaaaa gttagctgaa 360
tatggaccac aaggtagggc atttgtccat gagtgggctc atctacgatg gggagtattt 420
gacgagtaca ataatgatga gaaattctac ttatccaatg gaagaataca agcagtaaga 480
tgttcagcag gtattactgg tacaaatgta gtaaagaagt gtcanggagg cag 533

<210> 183
<211> 200
<212> DNA
<213> Homo sapiens

<400> 183
ctgctccttg tcatctccgg agctccagac ggtgcgcagg gcacgctcct ggttcctccg 60
tgccacccgg atcaggtaga ccatggaggc tcccaggaag aggatcaaca ccatcacgaa 120
cagccccgcc agaaccacca cctttgagcc aaggggcagg ggaggattct cctgctggcg 180
gaagcagcgg agcatgcagg 200

<210> 184
<211> 72
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(72)
<223> n = A,T,C or G

<400> 184
ctgagcanca caggccagga ggccacagtg taagcaataa cagatctgcc acatgcagaa 60
gcaaatatca gg 72

<210> 185
<211> 217
<212> DNA
<213> Homo sapiens

<400> 185
aaaaactctg gcttggatgt tacacagacc aacaacccaa acagcagcaa caacaacaca 60
aactcccccc acccccttct tcatcagccc caagattgtg aaaatgacag gaagtccagg 120
ttggctctgg catttatagc actgacatac attcagccca gagaagctct ggtgacaggc 180
tctctaaaca agtccctgtt cgggccccct ggtcagg 217

<210> 186
<211> 328
<212> DNA
<213> Homo sapiens

<400> 186
aaaatctcaa actaaaaatg ggaatcatac ttaaacatta gcattcccag gagagttgga 60

gcaagacctc tgtgcccact atcactcaac atttcattat ttaagtctta gcaagtgcaa 120
aacagaaaag atgcataaat attagagaag gaaagttatt ttttgcaaat ggcgtatttt 180
ttatctggaa ataccagaga atcaagtgag aaactttgaa gaataaaata attcagtaat 240
gatgcttcct accctaggtt aattaatata aagaagagtt aaattcctat ggcatatttc 300
tggtaacaaa aacccaggat cttgtttt 328

<210> 187
<211> 575
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(575)
<223> n = A,T,C or G

<400> 187
ctgagcagcc ctggatcttt gccgtactgt gactgggctc tttgccctat ttttccctct 60
gtctgtgccc ctggatggca ggctgaagtc agaggggctg tttcattctc agccccctca 120
gcagcactgg gggaagaaag cattgtcaca acaggttctt tctggccctc acccaacagc 180
ctgggcactt ggccctcctc ctccttgaca gccctccccc ttcctgcaaa ggacaggggc 240
gacaggggtt ggtgttggga ttggctcccg ctgcctgaca accacaagtt tatttggaag 300
gctagcggga agcccagcgg ctggcgtttc ccttgactaa ggaacagggt gcccatcaga 360
gtggggcggg cagctttggg aaggacacaa gaagcagtaa gagtgtaaag aggatgctgg 420
cctgggctca caccaatgcc acagtcagct tcctttctgc ccactgtgcc tctcaccttg 480
cgtggnttgg tgacagtctc accagtctct ctcagaggct acagatccag ctccccgatt 540
ccgtgaatca gctactccgc tatctgagag agctg 575

<210> 188
<211> 325
<212> DNA
<213> Homo sapiens

<400> 188
cctgtggccc tgcagaagag cccacgtgca aatccagctc ctcccagcag aacaacacag 60
tcctggtgga aggctgcttc tgtcctgagg gcaccatgaa ctacgctcct ggctttgatg 120
tctgcgtgaa gacctgcggc tgtgtgggac ctgacaatgt gcccagagag tttggggagc 180
acttcgagtt cgactgcaag aactgtgtct gcctggaggg tggaagtggc atcatctgcc 240
aacccaagag gtgcagccag aaacccgtta cccactgcgt ggaagacggc acctacctcg 300
ccacggaggt caaccctgcc gacac 325

<210> 189
<211> 222
<212> DNA
<213> Homo sapiens

<400> 189
ctgtatcact gaaaatttcc tgatattgta tgaaaaggaa aaagaaaaag atttcaaagt 60
gatccaggct acagtctcaa tgctgtaaaa ctacgtcggc gcccagccag gtgctgcaaa 120
ggagctcaga aaaatgaaaa gagccgaacc aggctagtgg aattccagat ctccctgctt 180
tagacacttc actttcatgt tattgtaaga ttttttttt tt 222

<210> 190
<211> 178
<212> DNA
<213> Homo sapiens

<400> 190

ctggaacaac tcagtgcaat gggatttttg aaccgtgaag caaacttgca agctctaata 60
gcaacaggag gtgatatcaa tgcagctatt gaaaggttac tgggctccca gccatcatag 120
cagcatttct gtatcttgaa aaaatgtaat ttatttttga taacggctct taaacttt 178

<210> 191
<211> 291
<212> DNA
<213> Homo sapiens

<400> 191
cctgagccac cgtgccttgc cacaaactat ttttaatcat taaaaaagaa aaaagaaaat 60
aaggcaggcc ggcttctcac attacatgct tttagaaaaa ggtctcatcc ttgaagcagc 120
tttgttatat gcagagcaca gtactggctt caaaaaatat ataaaggttc tgtgcactgg 180
cactgtttac atgtgaagaa ttgccatcaa cttctgtgaa aattagcaag ctggcacagt 240
ggctcacgcc tgtaatctca gcacctcagg aggccgaggt gggcggatca c 291

<210> 192
<211> 363
<212> DNA
<213> Homo sapiens

<400> 192
ctgccaaatg ggaagaatag aagaatttgc ccctaaaccc ctcctgtgtg ctgaccctgt 60
gctagacagt gctggagaca tagttggggg tggagaactg cccttatgga gcttgcagtc 120
cagtgaggtg gacagacctg tccccagaca gtgatggccc aaaatggtca ggactttaat 180
ggaggaggtg aggtgttgaa agcacaggca gagtggtcag ggctgaagtc ggagaagcac 240
agggactagg cccaatccag cctggaaagt cagggaggac ttcctagagg aagggacatc 300
gaactaagac ctgaactatg agaaataggc aggaagaagt tgtacctgac tcatttttct 360
cag 363

<210> 193
<211> 201
<212> DNA
<213> Homo sapiens

<400> 193
caggtactca tagtagctgt cggccctggc gcccagcgtg aatacgccca ggtgggtgaa 60
gaggccactg tgggtattga tgaacatggg caccagccca tccttcttcc cagacaggcc 120
gtggatgtgc tgtgtcacct tctccactgc ctcctgaaac ttcttatccc ctgtgagacg 180
ggagagctcc cggaactcca g 201

<210> 194
<211> 367
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(367)
<223> n = A,T,C or G

<400> 194
aaagttgaac taanattcta tcttggacaa ccagctatca ccaggctcgg taggtttgtc 60
gcctctacct ataaatcttc ccactatttt gctacataga cgggtgtgct cttttagctg 120
ttcttaggta gctcgtctgg tttcgggggt cttagctttg gctctccttg caaagttatt 180
tctagttaat tcattatgca gaaggtatag gggttagtcc ttgctatatt atgcttggtt 240
ataatttttc atctttccct tgcggtacta tatctattgc gccaggtttc aatttctatc 300
gcctatactt tatttgggta aatggtttgg ctaaggttgt ctggtagtaa ggtggagtgg 360

gtttggg 367

<210> 195
<211> 315
<212> DNA
<213> Homo sapiens

<400> 195
aaaaatattt acgtcttaca ggagctggat aatccaggtg caaaacgaat tctagagctt 60
gaccagttta aggggcagca gggacaaaaa cgtttccaag acatgatggg ccacggatct 120
gactactcac tcagtgaagt gctgtgggtc tgtgccaacc tctttagtga tgtccaattc 180
aagatgagtc ataagaggat catgctgttc accaatgaag acaaccccca tggcaatgac 240
agtgccaaag ccagccgggc caggaccaaa gccggtgatc tccgagatac aggcatcttc 300
cttgacttga tgcac 315

<210> 196
<211> 179
<212> DNA
<213> Homo sapiens

<400> 196
ctgtgcaaat gggcatgggg gtgcatggaa gggaggaaga gcaccgggcc cctgaacctg 60
ccccttttaa ggaggggggag gagccgtcag gccaggaagg ggaaatagtg caaggcagag 120
cccaggctgc aaaaggggtc cagcacccag cgaggaaggg ggggtgtccc ccaccccca 179

<210> 197
<211> 423
<212> DNA
<213> Homo sapiens

<400> 197
ttgttcttcg ggtcgctctt gaagtccccg ggcttcagga tgacgtgctg gttgtcgggc 60
cgtttgatga tacagtgcgt ctgttcacac ttcttacagc actccccggg ggcctccatg 120
agttcgaagc cagggctgca ggaggtgttg cagggcacgt gggtgcaggc gatgacgttg 180
agcagggtgt tgttgtccac cttgtccgtg cacacgcagt cctggcactt ggaggaataa 240
actggagaac cgggctggta ctcagcattc ccgtgaacac acacccccctt ggactcacac 300
cagtagaaag gacagcacct tccaggcacc atcttgctct tcacttcgaa tcccagcggg 360
cacacggagg gcttctcttt gcacaggctg gtgttgcact tgcagacggt aatgttgcag 420
cag 423

<210> 198
<211> 372
<212> DNA
<213> Homo sapiens

<400> 198
aaatgttctc atcagtttct tgccatgttg ttaactatac aacctggcta aagatgaata 60
tttttctact ggtattttaa tttttgacct aaatgtttaa gcattcggaa tgagaaaact 120
atacagattt gagaaatgat gctaaattta tagttttcag taacttaaaa agctaacatg 180
agagcatgcc aaaatttgct aagtcttaca aagatcaagg gctgtccgca acagggaaga 240
acagttttga aaatttatga actatcttat ttttaggtag gttttgaaag ctttttgtct 300
aagtgaattc ttatgccttg gtcagagtaa taactgaagg agttgcttat cttggctttc 360
gagtctgagt tt 372

<210> 199
<211> 502
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(502)
<223> n = A,T,C or G

<400> 199
ctgcagcctg ggactgaccg ggaggctctg attatttacc caccacaggt aggttgtgtt 60
ctgaatctca ggttcacagg ttaaggctac agcatcctca tcctccacgg ggttggagtt 120
gttgctggtg atgaagggtt tgggtggctc tgcatagact gtgatcgtcg tgactgtggt 180
cctattgagg ccagtgtctg agttatgggc ttggcacgta taggatccac tattattcac 240
agtgatgttg gggataaaga gctcttgggt ggattgctgg aaagtcccat tgacaaacca 300
agagtactgt gcaggtgggt tagaggctgc gtggcaggag aggttcagat tttcccctga 360
tctgtaagat gtgtttagag gggaaatggt gggggcatcc gggccataga ggacattcag 420
gatgactgaa tcactgcgcc tggcactcac tgggttctgg gtttcacatt tgtagcttgc 480
tgtgtcattt cttgngacat tg 502

<210> 200
<211> 609
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(609)
<223> n = A,T,C or G

<400> 200
ctgaaaagac agaagaatct aaggccgctg ctccagctcc tgtgtcggag gctgtgtgtc 60
ggacctccat gtgtagtata cagtcagcac cccctgagcc ggctaccttg aagggcacag 120
tgccagatga tgctgtagaa gccttggctg atagcctggg gaaaaaggaa gcagatccag 180
aagatggaaa acctgtgatg gataaagtca aggagaaggc caaagaagaa gaccgtgaaa 240
agcttggtga aaaagaagaa acaattcctc ctgattatag attagaagag gtcaaggata 300
aagatggaaa gccactcctg ccaaaagagt ctaaggaaca gcttccaccc atgagtgaag 360
acttccttct ggatgctttg tctgaggact tctctggtcc acaaaatgct tcatctctta 420
aatttgaaga tgctaaactt gctgctgcca tctctgaagt ggtttcccaa accccagctt 480
caacgaccca agctggagcc ccaccccgtg atacctcgca gagtgacaaa gacctcgatg 540
atgccttgga taaactctct gacagtctan gacaaangca gcctgaccca gatgagaaca 600
aaccaatgg 609

<210> 201
<211> 173
<212> DNA
<213> Homo sapiens

<400> 201
ctgcctcaga atgtcagaag ggtgccattt tcagttcttc tcctgtctga ctgtagggta 60
tagaacaggc agccctaagt gctgctttct gcgcaaatgt tttttgatta caaatctcta 120
actaggtttg aaatgtttta tagaataaga caatattctt ttcaacaaac ttt 173

<210> 202
<211> 182
<212> DNA
<213> Homo sapiens

<400> 202
gtgccacact ggcccttggt gttgttgcca aaccggtggt agggcagcct gacggagaag 60
gacaggccat tgtaggagac gaggacaccc agctcgggga tgtccaccac gtagttgatg 120

ccagactggt acacctccag cccgtacttc ttgtagggca gtgccaccgc ctgcctgttc 180
ac 182

<210> 203
<211> 106
<212> DNA
<213> Homo sapiens

<400> 203
gtgcctttga ggcagcaggg ttccacctga atgagcatct ctataacatg atcatccgac 60
gctactcaga tgaaagtggg aacatggatt ttgacaactt catcag 106

<210> 204
<211> 178
<212> DNA
<213> Homo sapiens

<400> 204
tctctccacc ctctgcagct catcgacagg accgagtccc taaaccgctc catagagaag 60
agtaacagtg tgaagaaatc ccagccagac ttgcccatct ccaagattga tcagtggctg 120
gaacaataca cccaggccat cgagaccgct ggccggaccc ccaagctagc ccgccagg 178

<210> 205
<211> 518
<212> DNA
<213> Homo sapiens

<400> 205
ctgcacaaac aatggaatgg tatttagtcc tctgataaca attcggttgt gaacatcccg 60
agctaggatg tgaagggctc cggtacaacc ttcaactatt tcttccatgc ggacccctc 120
cacaaattgc tgctgtgtcc cacccatgga cgtacggcgc tgggtatcct gatgtgcacg 180
aacaagcaac tgaactagtc gtggaatggc accctgctca cgcaaaggtg catgatttgc 240
gggacaaagg gcaagatttc gaatcaatcc aacagtagcc tttatcagag gccagtggga 300
tggtgggtgt aagagcttaa ccacaactgg tagtccatag tgaaggcgaa ctgcattctg 360
ggccatctct gcttcttggt gtcggctggt cagatgacga agagcacaga tggcaggctc 420
agtgatgtct tccctgtcac cagcccgaag gacagtacgc acaagagcct ctataccacc 480
cacttggcag accatcatct tgttcttata attattgc 518

<210> 206
<211> 367
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(367)
<223> n = A,T,C or G

<400> 206
aaagttgaac taagattcta tcttggacaa ccagctatca ccaggctcgg taggtttgtc 60
gcctctacct ataaatcttc ccactatttt gctacataga cgggtgtgct cttttagctg 120
ttcttaggta gctcgtctgg tttcgggggt cttagctttg gctctccttg caaagttatt 180
tctagttaat tcattatgca gaaggtatag gggttagtcc ttgctatatt atgcttggnt 240
ataatttttc atctttccct tgcggtacta tatctattgc gccaggtttc aatttctatc 300
gcctatactt tatttgggta aatggtttgg ctaaggttgt ctggtantaa ggtggagtgg 360
gtttggg 367

<210> 207

<211> 145
<212> DNA
<213> Homo sapiens

<400> 207
aaaaaaatta gattttagct ggagcttttg actaatgtaa agtaaatgcc aaactaccga 60
cttgataggg atgtttttgt aagttaattt tctaagactt tttcacatcc aaagtgatgc 120
tttgctttgg gttttaactg tttgg 145

<210> 208
<211> 193
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(193)
<223> n = A,T,C or G

<400> 208
ggccgagccc ccgatccccc actacgctgc catgtggcag cagcaggagc tactagaata 60
ttctcagcac aggaatgagg cttccttggt ttccatgtct gtaagggtta ctgatcactt 120
accttcttct ctttcagact tgaatctgnn gacatttctt tattgatatg gcaaattgct 180
tgcagatatt ttt 193

<210> 209
<211> 255
<212> DNA
<213> Homo sapiens

<400> 209
ccaggatgta gggccatcct taggggtggc ggcctctgcc cagaggggag aggttaatag 60
gtcaaggatg ggggccctga gggtatcaaa aggacaggga cagttcccag ttcccactat 120
ccagcatgct ctccactggc ccaggactcc atcgacacat cccaaagatt ccaatcaaag 180
tttgaggttg gctctcccaa actttcctct gcagagccat tcctgcaggc tccctcatgc 240
tggcaagcac cctcc 255

<210> 210
<211> 351
<212> DNA
<213> Homo sapiens

<400> 210
ctgaaagtgc ttctcagctc gccccatgta agttctcatt ccatgtaaat gacattttcc 60
agttacaact ggtactgaga ttttgcctct ctctttcctt actcatcctc ccaaatgtct 120
ttgtgggagc catatcagtg gataccaagc tctgtatcca tttgtccctt gccctccaca 180
atgtgtgaca tagaacaggg actttggccc tgggaaagca aaagctccca gtaaggaatc 240
ctgtgcccaa tgatgtaaaa caattccaaa catccaggaa tttttgtatc atagagcgaa 300
ttacttccta tcttttcatt agaggctatg aggacttcta attagtctca g 351

<210> 211
<211> 236
<212> DNA
<213> Homo sapiens

<400> 211
aaaaacccag aagatggggc agctcagaga ctggtttcct aatacacaag acctagcagg 60
aaatgatcaa gaaaatatta ggcatgcaga taggaacaac tctgatgata atcatttggc 120

ttcagaagat actagtgcca agcaaagtgg tgagccagac gcctgtcata ggcttcgtcc 180
tgagggtcca gcatgaggaa gatcaaatca aaccttgata ataaagtatg aggcag 236

<210> 212
<211> 135
<212> DNA
<213> Homo sapiens

<400> 212
ggacagggtt tcactgtgtt agccagaatg gtcttgatct cctggccttg tgatccgctc 60
acctcagcct cccaaagtgc tgggattaca ggcgtgagcc accacaccca gacttttttt 120
tttcttcttt ttttt 135

<210> 213
<211> 567
<212> DNA
<213> Homo sapiens

<400> 213
gtgcgcagcc ccgtcaccaa cattgctcgc accagcttct tccacgttaa gcggtccaac 60
atttggctgg cagcagtcac caagcagaat gtcaacgctg ccatggtctt cgaattcctc 120
tataagatgt gtgacgtgat ggctgcctac tttggcaaga tcagcgagga aaacatcaag 180
aacaattttg tgctcatata tgagctgctg gatgagattc tagactttgg ctacccacag 240
aattccgaga caggcgcgct gaaaaccttc atcacgcagc agggcatcaa gagtcagcat 300
cagacaaaag aagagcagtc acagatcacc agccaggtaa ctgggcagat tggctggcgg 360
cgagagggta tcaagtatcg tcggaatgag ctcttcctgg atgtgctgga gagtgtgaac 420
ctgctcatgt ccccacaagg gcaggtgctg agtgcccatg tgtcgggccg ggtggtgatg 480
aagagctacc tgagtggcat gcctgaatgc aagtttggga tgaatgacaa ggttgttatt 540
gaaaagcagg gcaaaggcac agacctc 567

<210> 214
<211> 470
<212> DNA
<213> Homo sapiens

<400> 214
aaaaaaccat ggaggtaaag aaaaagatta aactgcacat caggtagtca ataaaggcag 60
aatttactgt gcatgaaccc acatgagcaa agggtaagaa caggcaaatc agagaaaaat 120
ccaaacagaa acaaggacag caacaacaac aaacctcttt gaactcagac aaaaggcaat 180
taaactaaca agcaatacaa tgcaattttt agcctttcat attttcaagc attaaagagt 240
gctggagagg acgctggaac gggcgctttc attttggata gtaatcttgt aatatttctg 300
aaacatatgc ctacatagta tttctgggaa tccaacctat ataaataaaa gcaccagtat 360
gtattacagc agtgttattt tgaaaaaaaa taaaaaaagg aaataaaaga cgatcaataa 420
cgaaatggtt gaatgccttt ttggtacatc aacaagtact gtgtattcag 470

<210> 215
<211> 504
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(504)
<223> n = A,T,C or G

<400> 215
ttggtgcaca aaatactgtc atttgctcaa agctggctgc caaatgtttg gtgatgaagg 60
cannaatgaa tggctcaaaa cttgggagaa gagcaaaacc tgaaggggcc ctccagaaca 120

atgatgggct ttatgatcct gactgcgatg agagcgggct ctttaaggcc aagcagtgca 180
acggcacctc cacgtgctgg tgtgtgaaca ctgctggggt cagaagaaca gacaaggaca 240
ctgaaataac ctgctctgag cgagtgagaa cctactggat catcattgaa ctaaaacaca 300
aagcaagaga aaaaccttat gatagtaaaa gtttgcggac tgcacttcag aaggagatca 360
caacgcgtta tcaactggat ccaaaattta tcacgagtat tttgtatgag aataatgtta 420
tcactattga tctggntcaa aattcttctc aaaaaactca gaatgatgtg gacatagctg 480
atgtggctta ttattttgaa aaag 504

<210> 216
<211> 208
<212> DNA
<213> Homo sapiens

<400> 216
gtgttcccca ccttgggcat catgcaccac aacaaacagg ccactgagaa tgcagaggag 60
gaagtgaggc gaattctggg gctgctggat gcttacttga agacgaggac ttttctggtg 120
ggcgaacgag tgacattggc tgacatcaca gttgtctgca ccctgttgtg gctctataag 180
caggttctag agccttcttt ccgccagg 208

<210> 217
<211> 316
<212> DNA
<213> Homo sapiens

<400> 217
ccagctctgt ctcatacttg actctaaagt catcagcagc aagacgggca ttgtcaatct 60
gcagaacgat gcgggcattg tccacagtat ttgcgaagat ctgagccctc aggtcctcga 120
tgatcttgaa gtaatggctc cagtctctga cctggggtcc cttcttctcc aagtgctccc 180
ggattttgct ctccagcctc cggttctcgg tctccaggct cctcactctg tccaggtaag 240
aggccaggcg gtcgttcagg ctttgcatgg tctccttctc gttctggatg cctcccattc 300
ctgccagacc cccggc 316

<210> 218
<211> 327
<212> DNA
<213> Homo sapiens

<400> 218
ctgaagaaaa ctatcttccc agtcatgaga actcggggtg tcgattgagc agcctctcct 60
caacaatgga gatttttcag aaaagtgagt ctggtttaag ggacacaaat cccacagagt 120
ttcttggcac aatgtttcag taagcatgag tgagacccaa aaacgtccaa tgagagacca 180
acaggcccct gacctaagtc agagctagac tgcagaactc tcagcccagc tctgcagcag 240
acatctgcgt catcctccca gaaattcaac tcaggccaca cttctggcaa aagaagtttc 300
accctgactg tccttgtttt cagtcag 327

<210> 219
<211> 215
<212> DNA
<213> Homo sapiens

<400> 219
aaaatgtgaa tgccaggaat ggtgctgttg atatgaatat taggacaagg agcagacgtt 60
tcatcaggac taggtgtctc tggcggagtc tgtggaggaa taaacaaaga tactcgtgca 120
atgttggata tttctgattt cagatcgacc ttatcaacag cctgaatagc aatgaaaaga 180
tctgtgccat tttcaaaagt aatgttttct ggttt 215

<210> 220
<211> 344

<212> DNA
<213> Homo sapiens

<400> 220
gtggttgtag agcgactgca cataggtgaa gacacacttg gggtcaggct tcttgcccat 60
gatcatcatg tcgtccacct ccaccagggg cacacagtcc accagcatcc gtggggcccc 120
gagcaggggt taggactttt tggtttttac cagcccccttc tggaccagac agcggtagaa 180
ttcctggatg tacgtgtaca cgcacttcca gtcaggctct cgaagccgca ccatgtcctc 240
tgtatccagg agctgcgggc agtccgcatg ggtctccgca gatgagaagg ccacctcgaa 300
gttctggcgt cggttctgag ggctaagctg cccatagtcg aagg 344

<210> 221
<211> 262
<212> DNA
<213> Homo sapiens

<400> 221
gtgaatgggg cgggctgggc cacaggaagc ctcaccccag gacgggagaa cgtgggtgga 60
ggccacctgc cttggggagt ccagaggagc agcagccgtg agagctgtgg gggccggatt 120
ctggcagggc aggggcgttt gtcctgcagg gatgggggtc gctgtgtgtc cgcctctagg 180
ggctgtcagg cggtagctac tgcaccagcc ggtaggtgat gtacctgcca gccgtctcac 240
tgggccggat gatcttcacc ac 262

<210> 222
<211> 309
<212> DNA
<213> Homo sapiens

<400> 222
aaatggggtc attttacata cattattttg cattctgatt tcttttactt gacatgaatg 60
tattgtcaat aagtatcgat ctaattcaaa aatagctgca taacatatgg atgtatctaa 120
tttatccgaa cattctttat taaaagtcta tgagtggtct ggcgcagtgg ctcacgcctg 180
taatcccagc accccggggg gccgaggtgg gctgatcacc ctgaggtcag gagagaccgg 240
cctagccaac atgctgaaac cccgtctcta ctaataatac aaaaattagc caagtgtggt 300
ggcgcgcac 309

<210> 223
<211> 279
<212> DNA
<213> Homo sapiens

<400> 223
ctgcccccca cccttccctt cgatgacaac gtttgcaggc ttcaggggga ccagggaaca 60
aagctggggc ctggcagccc cactacgctg ccagccgggg agaacaagtc acaattacaa 120
attatcacaa caattagcgc ctgtacttgg gggatctgca aattgaggag gccccagctc 180
ctcattgtac acgggtctat ttggcagtga ccttgctctg gagacgatga tattccttca 240
gcctgaggga attgatgttg atgaacccgg tggcatcag 279

<210> 224
<211> 607
<212> DNA
<213> Homo sapiens

<400> 224
aaaatactgt catttgctca aagctggctg ccaaatgttt ggtgatgaag gcagaaatga 60
atggctcaaa acttgggaga agagcaaaac ctgaaggggc cctccagaac aatgatgggc 120
tttatgatcc tgactgcgat gagagcgggc tctttaaggc caagcagtgc aacggcacct 180
ccatgtgctg gtgtgtgaac actgctgggg tcagaagaac agacaaggac actgaaataa 240

cctgctctga gcgagtgaga acctactgga tcatcattga actaaaacac aaagcaagag 300
aaaaacctta tgatagtaaa agtttgcgga ctgcacttca gaaggagatc acaacgcgtt 360
atcaactgga tccaaaattt atcacgagta ttttgtatga gaataatgtt atcactattg 420
atctggttca aaattcttct caaaaaactc agaatgatgt ggacatagct gatgtggctt 480
attattttga aaaagatgtt aaaggtgaat ccttgtttca ttctaagaaa atggacctga 540
cagtaaatgg ggaacaactg gatctggatc ctggtcaaac tttaatttat tatgttgatg 600
aaaaagc 607

<210> 225
<211> 100
<212> DNA
<213> Homo sapiens

<400> 225
ctgtgtcttt agagctattg ccacattagc ctttgcactg tatagcgtct ggctttatgg 60
aacttaagtt taccaaatat aaaaagaaac ttctgctttt 100

<210> 226
<211> 260
<212> DNA
<213> Homo sapiens

<400> 226
ccactgataa ctcagtagcc atctgaatag tcatgcggtt taagaataca tccttgtata 60
atctgacata caaatttgtc atttcctgca catgcacacc attgttaaaa aaaaaagcca 120
gtaatagtgt ctggatcggt caggagcacg gcctctgagt cccctgtaat ttagttaagc 180
taaattaata cctcatacca aatggctcca ggaaaactgt cctgcaggtc agaagggagc 240
ccaagaagaa aagcacttgg 260

<210> 227
<211> 168
<212> DNA
<213> Homo sapiens

<400> 227
ctgtcaccat caagcccttg ctcttgggtt cattcttcgg gagccccagg ggcctatcta 60
tggggaaggg agtgggtgtg acttgggagc caatggaggg gtgggatggg tgagagaaaa 120
gggcagaatt cagatctgtt ttgtcttgga ttcttctgga actagagg 168

<210> 228
<211> 200
<212> DNA
<213> Homo sapiens

<400> 228
aaaaataact ggaatctgga ggccagtcaa acctttttgg acagatatct cctcgaaata 60
cttctttatt aaagtatttg aaggagatga cttcttagac ctttggcttg acattctggg 120
attcccagac tatattgttc tttcataagc aaataatctg ttatctttta gattcttcag 180
aataaataca tctttacttt 200

<210> 229
<211> 149
<212> DNA
<213> Homo sapiens

<400> 229
ccagaacacc gtgggctgtt aacttgcctt gagttggaag cggtttgcat ttacgcctgt 60
aaatgtattc attcttaatt tatgtaaggt ttttttttgta cgcaattctc gattctttac 120

ctgcccgggc ggccgctcga gccctatag 149

<210> 230
<211> 287
<212> DNA
<213> Homo sapiens

<400> 230
aaataaccaa gggcctccag aggcccctgt tcttcctttt gccctgtcta aaatccaatg 60
aaatatatta agtgttaaac tggtatgaag aggtggaact aaattctttg aaacacaagg 120
tggagtatca ctttttactt aaactttgag tcctttacat ttataactgc tattcaaaaa 180
aaattagaca aagacatcta gatttagatt aacgtgatca aagggattat tgtggatcat 240
taaaggaaac ttaacattaa gccttcatgt accaaatact aatattt 287

<210> 231
<211> 287
<212> DNA
<213> Homo sapiens

<400> 231
aaataaccaa gggcctccag aggcccctgt tcttcctttt gccctgtcta aaatccaatg 60
aaatatatta agtgttaaac tggtatgaag aggtggaact aaattctttg aaacacaagg 120
tggagtatca ctttttactt aaactttgag tcctttacat ttataactgc tattcaaaaa 180
aaattagaca aagacatcta gatttagatt aacgtgatca aagggattat tgtggatcat 240
taaaggaaac ttaacattaa gccttcatgt accaaatact aatattt 287

<210> 232
<211> 222
<212> DNA
<213> Homo sapiens

<400> 232
ctggggccac tgtcggcatc atgattggag tgctggttgg ggtggctctg atatagcagc 60
cctggtgtat tttcgatatt tcaggaagac tggcagattg gaccagaccc tgaattcttc 120
tagctcctcc aatcccattt tatcccatgg aaccactaaa aacaaggtct gctctgctcc 180
tgaagcccta tatgctggag atggacaact caatgaaaat tt 222

<210> 233
<211> 536
<212> DNA
<213> Homo sapiens

<400> 233
ccaacatggt aaaaccccat ctctactaaa aatacaaaaa ttagctgagc gtggtggcgg 60
gcacctgtaa tcccagctac tcaggagact gaggcaggag aatcatttga acccgggagg 120
cagatgttgc cagtgagctg agatcacgcc attgcactcc agcctgggcg acaagagcaa 180
aactcaaaaa aaaaaaaaaa aaaggaaaaa caaaacccag aatgcagaaa ctgcattggt 240
attttgagct agttacaggc agtaacactt ctactaggaa ggaaaatata taatttatcc 300
aaaaatagtt ctaagatata gaaaaacatc tcaatcctct agggccaaga cctgcccctt 360
atttacattc acaaagccat taagtgaacc cagaactgac cagcagacag tgaggcaggg 420
cctgctctgt ggcacatgcc agtcacctac tgcgtaagtg gacaaagaga aagcagaagg 480
taatggagag agttttcatt tgcttattta gtgaaaaaca gaggaaaatc actcga 536

<210> 234
<211> 562
<212> DNA
<213> Homo sapiens

<400> 234
ttggtgcaca aaatactgtc atttgctcaa agctggctgc caaatgtttg gtgatgaagg 60
cagaaatgaa tggctcaaaa cttgggagaa gagcaaaacc tgaaggggcc ctccagaaca 120
atgatgggct ttatgatcct gactgcgatg agagcgggct ctttaaggcc aagcagtgca 180
acggcacctc cacgtgctgg tgtgtgaaca ctgctggggt cagaagaaca gacaaggaca 240
ctgaaataac ctgctctgag cgagtgagaa cctactggat catcattgaa ctaaaacaca 300
aagcaagaga aaaaccttat gatagtaaaa gtttgcggac tgcacttcag aaggagatca 360
caacgcgtta tcaactggat ccaaaattta tcacgagtat tttgtatgag aataatgtta 420
tcactattga tctggttcaa aattcttctc aaaaaactca gaatgatgtg gacatagctg 480
atgtggctta ttattttgaa aaagatgtta aaggtgaatc cttgtttcat ttctaagaaa 540
atggacctga cagtaaatgg gg 562

<210> 235
<211> 313
<212> DNA
<213> Homo sapiens

<400> 235
ctggtgtgat gggaacttca gtggctacat ctgctagcaa aattattcct caaggggccg 60
atagcacaat gcttgccacg aaaaccgtga aacatggtgc acctagtgac ctgagatccg 120
aactgggctc atcaagggca gcggcactgc agaggtggag ctgaagaagg gagccactct 180
caaaatcacg ctggataacg cctacatgga aaagtgtgac gagaacatcc tgtggctgga 240
ctacaagaac atctgcaagg tggtggaagt gggcagcaag atctacgtgg atgatgggct 300
tatttctctc cag 313

<210> 236
<211> 172
<212> DNA
<213> Homo sapiens

<400> 236
gtgcgcgcca ccaagcccag ctcatttttg tatttttagt agagatgggg tttcacgatg 60
ttggctagga tggtctcgat ctctggtcag agtcttttct gtaaatatcc ttggtaaaga 120
agcaatttta gactgtagct gttgcaaatg ctttaaggaa gaagcaaaac aa 172

<210> 237
<211> 454
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(454)
<223> n = A,T,C or G

<400> 237
ctggggcatg tttatttcct cctgggcctg gcaggctggg agcagaatat anacaaaggc 60
actggggcac ctgggtctgg cacggtctgg ancttggccg nctgggtagc aaccgtnaag 120
ggtgtgccag ggcgngcang gactggagtn atnctnccag aactgagaga gggccctcgg 180
ggcatggggg catcacaagt gctaggcttg gcacaggtac aggggagagg ttacggagtg 240
ggtgtgtgca gggcctggtg ggaatgggga gacccgtgga cagagcttgt tagagtgtcc 300
tagagccagg gggaactcca ngcagggcaa attgggccct ggatgttgag aagctgggta 360
acaagtactg agagaacaaa agcttgtggg tcagcangcc ccacaaagat gtgactgcag 420
acaggatcgg ccctgggaga gaccgaggct ccag 454

<210> 238
<211> 331
<212> DNA

<213> Homo sapiens

<400> 238
aaaatactaa gtcatcttac gtttccattt tattaacggg atgttgcaat cgtttgtaaa 60
ctaataaact tataaagtga ttggcacaaa gactccttga gcaaaagctg tgcagttaag 120
tacaaaaaga tacttaattt ggagactctt acagtaattt ttgccatgtc aaaacaatgg 180
cttttacatt gaaagattaa tagaaactct acatatgtta atttttttat agaacctgac 240
tcaaatcaag gtactctcca ttttattgcc ttacctgaat cagtcctttt tggttggtaa 300
tagatttttt tatacaccca cgtttgattt a 331

<210> 239
<211> 353
<212> DNA
<213> Homo sapiens

<400> 239
gtggtgcagt gggtgagctt tgctgattcc gatatagtgc ccccagccag tacctgggtg 60
ttccccacct tgggcatcat gcaccacaac aaacaggcca ctgagaatgc aaaggaggaa 120
gtgaggcgaa ttctggggct gctggatgct tacttgaaga cgaggacttt tctggtgggc 180
gaacgagtga cattggctga catcacagtt gtctgcaccc tgttgtggct ctataagcag 240
gttctagagc cttctttccg ccaggccatg tggcaagcga caggcacaaa acaattttcc 300
aagtcaatag gaaaaacctc agagctgaaa tctttatatg ctgtactaca cag 353

<210> 240
<211> 356
<212> DNA
<213> Homo sapiens

<400> 240
aaagaactac aagccctcag actctaccta gcggcggcag tgcccgcggt ctgctatacg 60
atgtactcca ttcggtttaa gctctgggca ctttccaagt ctctgttgtc ctgtttgttt 120
gtccagtttg ggtgttttgt cggcgtgccg ttgggggggct tctcttctct gtctaccagc 180
gtgtacgccg gctgcttggc aaaccgggct ttctgctggt gtttgtccat gtcgtcctct 240
tctacttcag aattgtgtgt ccttatttta gacattttgg agttcttgtt ctcgtaatcc 300
ttgatgggga ccgtgttggc cccatgtttc tcaatggggt ttttgatctg gttcag 356

<210> 241
<211> 425
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(425)
<223> n = A,T,C or G

<400> 241
gtggagtcca caatcttggg aaggcgtcta tttcttcttg tgtgtactca tctagacgtt 60
taggtatttt tcgtggttga ggaagctnct ctactaaatt cttaagaata tcttctggaa 120
tatactcatc tggaaaaaga tgcaaccttt ccatcattgt tcttctgtga aggttttttg 180
gcagcatgcc ataaatagct agttttacaa ttgccactgg atccctcagg tgaagctgan 240
cagctgttac ttgtctaaat ccacctgggt agtcagtgca tggtacacag gtttatgtaa 300
tccctgaagt cttatagatg ccatagcngc aagtttgcca ggtggctgca ttttcccatc 360
taagagatac catattctag caaaagtggc ccattgntgg ggcgccctag agaaactcga 420
cacct 425

<210> 242
<211> 101

<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(101)
<223> n = A,T,C or G

<400> 242
ctgtctggga ccacacggnc accggcctnc tgtgagcgga tccactcact gtctcgccag 60
tcctgtgagt tgttgctgcc acagcagtgg aactnctgct g 101

<210> 243
<211> 284
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(284)
<223> n = A,T,C or G

<400> 243
gtgatctgcc agccttggct tcccaaagct ctaggattac aggcgtgagc cactgcaccc 60
agccaagagg tagtttctta aaggttantt gcagcanaat ctgaaaccat aaaaaggaaa 120
ttttcatgct ctgttacatt aaaattggnt ggcccatctt gaactttgaa tggactgctt 180
acccatgcat gattctgtaa catcatggnt ggttatttgg aaaatattgg ttcactgaat 240
tgtgaanatc ttccaaatat tgaaacattt cantatgtat tttt 284

<210> 244
<211> 266
<212> DNA
<213> Homo sapiens

<400> 244
ctgagctcac catagtctaa tagaaaacaa ccgaaaccaa ataattcaag cactgcttat 60
tacaatttta ctgggtctct attttaccct cctacaagcc tcagagtact tcgagtctcc 120
cttcaccatt tccgacggca tctacggctc aacatttttt gtagccacag gcttccacgg 180
acttcacgtc attattggct caactttcct cactatctgc ttcatccgcc aactaatatt 240
tcactttaca gccaaagtga tgtttg 266

<210> 245
<211> 432
<212> DNA
<213> Homo sapiens

<400> 245
ctgctcagga tgctgaggca ggagaatcac ttgaacctag gaggcatagg ttgcagtgaa 60
ctgagatcgc gccactgcac tccagcctgg gtgacagagt gagactctgt gatacagttt 120
ccatctggct ctcagtctta ggacctgtgg aactagccac cctgttgtga ggcagccaga 180
ccacatgaag aggccatgtg taggtgttct ggctgatagc cccagccatc agtcagtatt 240
aaccactaga caagcaagtg gacaagcctt cagatgatgc cagctccagc ctttgagctg 300
ccccagctaa taccacatgg agcagacact atccctgcag agtcactgtt ttaagctact 360
aagtgttaag tgtttttga ggtagcaata gataactggg aaattctcaa acctattttg 420
catcctcttt tt 432

<210> 246
<211> 367

<212> DNA
<213> Homo sapiens

<400> 246
ctgctgtcct aaacccaact tctctgtgat gcccggattt cttgattttg atccagtagc 60
tgctcatttt cctgcctttt acatttagga gattcaagct ctgtcatttc ctctagctgc 120
ccctgaagtc cgtccttcct gcagggccca actccacgta gagtgagtgc agccacacag 180
cagtaaccag atagagcagc ctcccctgca gacatgagca aagaagggat ccagagagcc 240
aaggctgtat catagattct tgtggggtca aaggggcagt cagtatgtcc cggcccctca 300
tccagtggta ccagaggatc cagcagtcct gggtggcagt cagcaataag gcggcggcca 360
ccgttgg 367

<210> 247
<211> 105
<212> DNA
<213> Homo sapiens

<400> 247
aaatgacaga agcgtatatg aattcaagaa aatttaagct gcaaaaatgt atttgctata 60
aaatgagaag tctcactgat agaggttctt tattgctcat ttttt 105

<210> 248
<211> 538
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(538)
<223> n = A,T,C or G

<400> 248
ccattgaaga gttgatttga acgatgacgt cccatggtgg gtgaaaatct atcctggata 60
actcctggac cagtaccaat tccgctacct ggcatttgtc caaacatatc agcaagtcct 120
ccaagtgggt ccctatccat tttcatcctg ggtggcatga acggtccctc cagaaagaag 180
tcacttctca tcccttgagc cataggagca ggaataaaca cccctagatc ttttactgca 240
tcttgacgaa tttgattgat cgtctttggt ccattgtcaa gaaaagcctt gcgaggaacc 300
caatggtgtt ctcgcaactc tacggtatcc tgcagcagga aacgaatcct tgctggcaat 360
tccttactta acatcaagga gcacattcgg gcaaagtact gatccattaa ggacttggct 420
cgttcatggt ctaatctang tcccactgtc ctcattatct gacagaggca ctccaaatcc 480
tctcccatat ctttgagttg gactctcttc ttcttttcca aaagtgtttt gatgcact 538

<210> 249
<211> 557
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(557)
<223> n = A,T,C or G

<400> 249
ctgacacaga ggtggtggaa gactccttga ggcagcgtaa aagtcagcat gctgacaagg 60
gactgtagat ttaatgatgc gttttcaaga atacacacca aaacaatatg tcagcttccc 120
tttggcctgc agtttgtacc aaatccttaa tttttcttga atgagcaagc ttctcttaaa 180
agatgctctc tagtcatttg gtctcatggc agtaagcctc atgtatacta aggagagtct 240
tccaggtgtg acaatcagga tatagaaaaa caaacgtagt gttgggatct gtttggagac 300

tgggatggga acaagttcat ttacttaggg gtcagagagt ctcgaccaga ggaggccatt 360
cccagtccta atcagcacct tccagagaca aggctgcagg ccctgtgaaa tgaaagccaa 420
gcaggagcct tggctcctga gcatccccaa agtgtaacgt agaagccttg catccttttc 480
ttgngtaaag tatttatttt tgtcaaattg caggaaacat caggcaccac agtgcatgaa 540
aaatctttca cagctag 557

<210> 250
<211> 465
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(465)
<223> n = A,T,C or G

<400> 250
ctgtgtcaaa agaaaaacag agtctacaaa tcttactaaa tacaaccatt ggcaaacagt 60
catatttaat ggttcagaaa aggagagaca gtaaaggaat ttttcttttc aactttgctc 120
cggaggaggc tgtctggagc ccattccttt ggcctcgact tttcagtgtt ataactgtcc 180
ttgaagtaag ctggctaagc agaggaaaac ttgttcttgt tttcttttta acccttaccc 240
cctgccacat aatcacatct ttacacttct ttttttttt ttttaanatg ggaagtcgga 300
gtctcgctct gtcgcccagg ctggagtgca gtggtgccat ctcggctcat tgtaacctct 360
gcctcccagg ttcaagtgat tcaagccatt cttgtgcctc tgcctcctga gtagctggga 420
ttataggcnc acaccaccac acctggcgaa ttttttttt ttttt 465

<210> 251
<211> 429
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(429)
<223> n = A,T,C or G

<400> 251
ctgtagataa caatagcaaa gtgaagacag catctaagct gtgttcagcc ctgactcttt 60
ctggtcttgt ggaagtgaaa gagctgcagc gggagcccct aacccctgag gaagtacagt 120
ctgttcgaga acaccttggt catgaaagtg acaacctgct gtttgttcag atcacaggca 180
aaaaaccaaa ctttgaagtg ggttcttcta ggcagcttaa gctttccatc accaanaagt 240
cttctccttc agtgaaacct gctgnggacc ctgctgntgc caagctgtgg accctctcag 300
ccaacgatat ggaggacgac agcatggatc tcattgactc agatgagctg ctggatccag 360
aagatttgaa gaagccagat ccagcttccc tgcgggctgc ttcttgtggg gaagggaaaa 420
agaggaagg 429

<210> 252
<211> 559
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(559)
<223> n = A,T,C or G

<400> 252
aactactgaa ttccaagctg cctcggnggc aggagacctg tgttgatgcc atcaaagtgc 60

cagagaaaat catgaatatg atcgaagaaa taaagacccc agcctctacc cccgtgtctg 120
gaactcctca nncttcaccc atgatcgaga gaagcaatgt ggttaggaaa gattacgaca 180
ccctttctaa atgctcacca aagatgcccc ccgctccttc aggcagagca tataccagtc 240
ccttgatcga tatgtttaat aacccagcca cggctgcccc gaattcacaa agggtaaata 300
attcaacagg tacttccgaa gatcccagtt tacagcgatc agtttcggtt gcaacgggac 360
tgaacatgat gaagaagcag aaagtgaaga ccatcttccc gcacactgcg ggctccaaca 420
agaccttact cagctttgca cagggagatg tcatcacgct gctcatcccc gaggagaagg 480
atggctggct ctatggagaa cacgacgtgt ccaaggcgag gggttggttc ccgtcgtcgt 540
acacgaagtt gctggaaga 559

<210> 253
<211> 181
<212> DNA
<213> Homo sapiens

<400> 253
ccaggatggc ctcgcgctcc ctatcggcgt ccggcagggt ggacttgaac tggtcatggg 60
ctgagatcag gccctcaatc tcctcgatgg tatggacgat gaacatgtcc tggaggtcct 120
ccatggcgct ctccatccag ttgttgaagg gggccgcgcg cttggcgtat tccaggtgca 180
g 181

<210> 254
<211> 137
<212> DNA
<213> Homo sapiens

<400> 254
ccaccacctc ggctccaaag gtaaaagaga cgccacggtc gttctcgccc cagccctgca 60
cgtccttgtc agggtcagac cacagcaggt cacacagcag gccctggtca ggcacatctg 120
tgggccgcat gatccgc 137

<210> 255
<211> 193
<212> DNA
<213> Homo sapiens

<400> 255
ctgcggccca tgatgtcaga gctggaagag agggcacgtc agcagagggg ccacctccat 60
ttgctggaga caagcataga tgggattctg gctgatgtga agaacttgga gaacattagg 120
gacaacctgc ccccaggctg ctacaatacc caggctcttg agcaacagtg aagctgccat 180
aaatatttct caa 193

<210> 256
<211> 532
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(532)
<223> n = A,T,C or G

<400> 256
ctgcgtgagc tcactgtcag acaagatgga agaagaaggg ctggagtgtc caaactcttc 60
ctctgaaaaa cgctattttc ctgaatccct ggattccagc gatggggatg aggaagaggt 120
tttggcctgt gaggatttgg aacttaaccc ctttgatgga ttgccatatt catcacgtta 180
ttataaactt ctgaaagaaa gagaagatct tcctatatgg aaagaaaaat actcctttat 240
ggagaacctg cttcaaaatc aaatcgtgat tgtttcagga gatgctaaat gtggtaagag 300

cgctcaggtt cctcagtggt gtgctgaata ttgtctttcc atccactacc agcacggggg 360
cgtgatatgc acacaggtcc acaagcagac tgtggtccag cccgccctgc gggtggcgga 420
tgaaaaatgg atgttaacat tggtcatgan ggtttggcta cnntgatccc tttcggagaa 480
ctgctgtacc aangaaacaa tcctgangna ttgnactgat gatatgctnc aa 532

<210> 257
<211> 300
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(300)
<223> n = A,T,C or G

<400> 257
ctaataatta ggctgtgggt ggttgtgttg attcaaatta tgtgtttttt ggaaagtcat 60
gtcagtggta gtaatataat tgttgggacg attagtttta ncattggagt aggtttaggt 120
tatgtacgta ntctaggcca tatgtgttgg anattganac tantanggct aggcccaccn 180
ntgcttcgca nncggcaaag actagtatgg caataggcac aatattggct aanagggagt 240
gggtgttgan ggttatnana ntagctntan tgaacancga tagcattatt ccttctaggc 300

<210> 258
<211> 308
<212> DNA
<213> Homo sapiens

<400> 258
ggccagccca ccctcctggg gctgacatga gccattccct gtgatgttca ctctcctccc 60
aaagcaaacc acagccaagc ctgtctgagc tgggagtccc cttccccagc agagctccca 120
gtccctgcat acccagcggg gtggcgactc gggaagagct gagctggaga cggctctaga 180
ccaagtccgg ccaaccaggc ggttctgtaa tcctctccca gggcccatgg aagttaggct 240
tccatcaggc gcacttttcc caccaggggc tctgggagga cgtgtcttct aaagtgttcc 300
gtgctcac 308

<210> 259
<211> 344
<212> DNA
<213> Homo sapiens

<400> 259
cctaacctcc atgtgcaggc ccagtttggc actctttctg actattttga tgccctgtac 60
aagaggacag gggtggagcc aggggcccgg cctccagggt ttcctgtgct gagcggggat 120
ttcttctcct atgcggaccg ggaggatcat tactggacag gctattacac ttcccggccc 180
ttctacaaga gcttagaccg agtccctaaa ccgctccata gagaagagta acagtgtgaa 240
gaaatcccag ccagacttgc ccatctccaa gattgatcag tggctggaac aatacaccca 300
ggccatcgag accgctggcc ggacccccaa gctagcccgc cagg 344

<210> 260
<211> 416
<212> DNA
<213> Homo sapiens

<400> 260
cctacagact tatttcttct tggacacacc cacggcgcgg ccacggcggc cagtggtctt 60
ggtgtgctgg cctcggacac gaaggcccca gaagtgacgc agccctctat gggcccgaat 120
cttcttcagt cgctccaggt cttcacggag cttgttgtcc agaccattgg ctaggacctg 180
gctgtatttt ccatccttta catccttctg tctgttcaag aaccagtctg ggatcttgta 240

ctggcgtgga ttctgcataa tggtgatcac acgttccacc tcatcctcag tgagttctcc 300
cgccctcttg gtgaggtcaa tgtctgcttt cctcaacacc acatgagcat atcttcggcc 360
cacaccctta atggcagtga tggcaaaggc tattttccgc cgcccatcga tgttgg 416

<210> 261
<211> 189
<212> DNA
<213> Homo sapiens

<400> 261
aaaacaagtg tgatgccata tcaagtccat gttattctct cacagtgtac tctataagag 60
gtgtgggtgt ctgtttggtc aggatgttag aaagtgctga taagtagcat gatcagtgta 120
tgcgaaaagg tttttaggaa gtatggcaaa aatgttgtat tggctatgat ggtgacatga 180
tatagtcag 189

<210> 262
<211> 219
<212> DNA
<213> Homo sapiens

<400> 262
ctgtatcaac acatcccagc atcgtcaagg agacactgcc tctgctgctg cagcatctct 60
ggcaagtgaa cagagggaat atggttgcac aatccagtga cgttattgct gtctgtcaga 120
gcctcagaca gatggcagaa aaatgtcagc aggaccctga gagttgctgg tatttccacc 180
agacagctat accttgcctg cttgccttgg ctgtgcagg 219

<210> 263
<211> 193
<212> DNA
<213> Homo sapiens

<400> 263
aaagtttgtg ctataaaatt gtgcaaatat gttaaggatt gagacccacc aatgcactac 60
tgtaatattt cgcttcctaa atttcttcca cctacagata atagacaaca agtctgagaa 120
actaaggcta accaaactta gatataaatc ctaccaataa aatttttcag ttttaagttt 180
tacagtttga ttt 193

<210> 264
<211> 605
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(605)
<223> n = A,T,C or G

<400> 264
tcaggaggca gcgctctcgg gacgtctcca ccatggcctg ggctctgcta ttcctcaccc 60
tcctcactca gggcacaggg tcctgggccc agtctgccct gactcagcct gcctccgtgt 120
ctgggtctcc tggacagtcg atcaccatct cctgcactgg aaccagcagt gacgttggtg 180
gttataacca tgtctcctgg taccaacaac acccaggcaa agcccccaaa ctcatgattt 240
atgatgtcac tagtcggccc tcaggggttt ctaatcgctt ctctggctcc aagtctggca 300
acacggcctc cctgaccatc tctgggctcc aggctgagga cgaggctgat tattactgca 360
gctcatatac tagcatcatc actgtggtat tcggcggagg gaccaaggtg accgtcctag 420
gtcagcccaa ggctgccccc tcggtcactc tgttcccgcc ctcctctgag gagcttcaag 480
ccaacaaggc cacactggtg tgtctcataa gtgacttcta cccgggagcc gtgacagtgg 540
cctggaggna gatagcagcc ccgtcaaggc gggagtggag accaccacac cctncaaaca 600

aagca 605

<210> 265
<211> 593
<212> DNA
<213> Homo sapiens

<400> 265
ctgttactga agaggaaccc tgtcatgttc ttccaacact tcattgaatg tatttttcac 60
tttaataact atgagaagca tgagaagtac aacaagttcc cccagtcaga gagagaagcg 120
gctgttttca ttgaagggaa agtcaaacaa agagagacga atgaaaatct acaaatttct 180
tctagagcac ttcacagatg aacagcgatt caacatcact tccaaaatct gccttagtat 240
tttggcgtgc tttgctgatg gcatcctacc cctggacctg gacgccagtg agttactctc 300
agacacgttt gaggtcctca gctcaaagga gatcaagctt ttggcaatga gatctaaacc 360
agacaaagac ctccttatgg aagaagatga catggccttg gcaaatgtag tcatgcagga 420
agctcagaag aagctcatct cacaagttca gaagaggaat ttcatagaaa atattattcc 480
aattatcatc tccctgaaga ctgtgctgga gaaaaataag atcccagctt tgcgggaact 540
catgcactat ctcagggagg tgatgcagga ttaccgagat gagctcaagg act 593

<210> 266
<211> 461
<212> DNA
<213> Homo sapiens

<400> 266
ctgatccagc gcttccagga tgagcacgag gactggtcca ttatcatctt caccaacacg 60
tgcaagacct gccagattct gtgcatgatg ctgcgcaaat tcagcttccc caccgtggct 120
ctgcactcca tgatgaagca gaaagaacgc tttgccgccc tagccaagtt caagtccagc 180
atctaccgga tcctgatcgc aacagacgtg gcctcccggg gcctggacat ccctacggta 240
caggtggtca tcaaccacaa caccccggg ctccccaaga cctccaagga tccttgcagg 300
agttgaagga gagggctctg agccgataca acctcgtgcg gggccagggt ccagagaggc 360
tggtgtctgg ctccgacgac ttcaccttat tcctgtggtc cccagcagag gacaaaaagc 420
ctctcactcg gatgacagga caccaagctc tcatcaacca g 461

<210> 267
<211> 489
<212> DNA
<213> Homo sapiens

<400> 267
ccttcgagaa gatccctagt gagactttga accgtatcct gggcgaccca gaagccctga 60
gagacctgct gaacaaccac atcttgaagt cagctatgtg tggtgaagcc atcgttgcgg 120
ggctgtctgt agagaccctg gagggcacga cactggaggt gggctgcagc ggggacatgc 180
tcactatcaa cgggaaggcg atcatctcca ataaagacat cctagccacc aacggggtga 240
tccactacat tgatgagcta ctcatcccag actcagccaa gacactattt gaattggctg 300
cagagtctga tgtgtccaca gccattgacc ttttcagaca agccggcctc ggcaatcatc 360
tctctggaag tgagcggttg accctcctgg ctcccctgaa ttctgtattc aaagatggaa 420
cccctccaat tgatgcccat acaaggaatt tgcttcggaa ccacataatt aaagaccaga 480
cctgcccgg 489

<210> 268
<211> 242
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(242)

<223> n = A,T,C or G

<400> 268
aaataaaaaa gctatatnnn aaagtaacct aggagggcca ggcacagtgg ctcatgccta 60
ttacctcanc actttgggag gcagaggcca gaggactgct cgagcccagg agtttgagac 120
cagcctgggc aacatgggga gaccccatct cttcagaaaa caaaaaggtc agccaggcat 180
agtggcacac ttggtggtcc cagctattca ngangctgag gtggtggatc acacctcggc 240
cg 242

<210> 269
<211> 320
<212> DNA
<213> Homo sapiens

<400> 269
aaagaattta ttaagcctgt tataccacac agtatgtttt atacactgac atacaactcc 60
ctaataagat aaagcaaaga caaaaaagtt tatcttatta gaaacaagat acaccaccac 120
ttattgtctt cagacattat tgcactttaa ctttcttaat ttgacaaagc attcaagaaa 180
catctgcaga ctagttttaa cagacaaata acacctgtaa gcagacatga ctgtcctaaa 240
ttgtttatta agtatgaatt ttacaaactt tacttatatt agcggtaacg gtggagctgg 300
agagtattgc gccttctcca 320

<210> 270
<211> 400
<212> DNA
<213> Homo sapiens

<400> 270
aaatccgcgc cctgcacacg caattcattt agacctttc gtgaatcttc tccactttca 60
caaacaacct atccagatca ttcctcaggt catctagtaa acccttggct gattccagat 120
tgttctcgtt ggtttctatt ttgaccgagt atgcaaccaa actgtccaca gcagtcctga 180
gcattttcaa gtccgcctcc acttggctga ctgaggcttt caggttgtct agagaagaaa 240
gtctgtccag gaagtcctga ggaggcagac gggcggcctg ggcttggtcc tgactgagca 300
gcgtgtgcac ctgctcctgc accttctgga gtgattccac ggtgctgggg agctcgccca 360
cactcctctt cagctcctcc acgtcaccgt agagcaccag 400

<210> 271
<211> 536
<212> DNA
<213> Homo sapiens

<400> 271
aaaaaagcaa cttccagggt tgtcattgta caggttttgc ccagtctcct atagcatggt 60
atagtgataa ctgatttttt ataacaatga ctcagaggca ttgaagatcc ataactatct 120
tctgaattat cacagaaaga agaaagttag aagagtttaa tgttaagtgt attaaaaatc 180
atattctaat tcttttaatt tggttatctg agtatgataa tataggagag ctcagataac 240
aagaaaaggc aattggttag aacactccat tcccacagga tgtgcattaa cagacttttt 300
actgcatatg tctttatata gtttgcaaac taattcaacc attttacaca gcattaattt 360
ttttttaact gggttgacat tgggctgaaa catttgctta tcatcttata attatttttt 420
cctgttcttt aatggatttt acccccatct gacatagtgt ttggacttta gtgtatgtga 480
cacttcaaga tcatctctgc ccattctgat gatagttaca atgaggttac ccatgg 536

<210> 272
<211> 424
<212> DNA
<213> Homo sapiens

<400> 272

aaaatgagca gtgttcgtct tcgtaaagaa attaagagaa gaggcaagga ccccacagaa 60
cacatacctg aaataattct gaataatttt acaacacggc tgggtcattc aattggacgt 120
atgtttgcat ctctctttcc tcataatcct caatttatcg gaaggcaggt tgccacattc 180
cacaatcaac gggattacat attcttcaga tttcacagat acatattcag gagtgaaaag 240
aaagtgggaa ttcaggaact tggaccacgt tttaccttaa aattaaggtc tcttcagaaa 300
ggaacctttg attctaaata tggagagtat gaatgggtcc ataagccccg ggaaatggat 360
acaagtagaa gaaaattcca tttataaagt actgagagaa tgatattgga ttttgctgaa 420
cagg 424

<210> 273
<211> 232
<212> DNA
<213> Homo sapiens

<400> 273
aaataaaaaa gctatatgta aaagtaacct aggagggcca ggcacagtgg ctcatgccta 60
ttacctcagc actttgggag gcagaggcca gaggactgct cgagcccagg agtttgagac 120
cagcctgggc aacatgggga gaccccatct cttcagaaaa caaaaaggtc agccaggcat 180
agtggcacac ttggtggtcc cagctattca ggaggctgag gtggtggatc ac 232

<210> 274
<211> 112
<212> DNA
<213> Homo sapiens

<400> 274
gtgatccacc cgcctcggcc tcccaaagtg ctgggattac aggcttgagc ccccgcgccc 60
agccatcaaa atgcttttta tttctgcata tgttgaatac tttttacaat tt 112

<210> 275
<211> 468
<212> DNA
<213> Homo sapiens

<400> 275
aaaaagcagc ctgggcaaga gaagtgggtg ggtttaggag aatccctttc gaaaaattca 60
gagcattatt attaatcctt cttaaattaa atgcagggcc aagcatgctg cacgtggaat 120
ctggacaatt ttttgataaa ctttaaggct gctaaataat ttacagaaac tgtgaatgca 180
ttttcatttt acgaggcaaa agagaaaata ttcaagattg catagcaatt ttatttttg 240
aaatggttat cctaaagaat ttccttaaat tcagattttg caaaattcct actctccaag 300
tcatcaagtg aacactaaaa gcaactttac tcgtgaatac agtggactct ttacgaggca 360
tgcatttttc ataaatctag gccaaagtga actaattgag atttaattct aaattcatcc 420
tgtgatttct gcatataata ttggtataaa accagtaaaa atactttt 468

<210> 276
<211> 461
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(461)
<223> n = A,T,C or G

<400> 276
ccttctcctg agctgaaagt tctttggcag atgagcaaga aactgaaagc tgatgtacct 60
gactggctct gtaagatcag aaaactgtat ccagaataag ccctatggat taacccctga 120
gtacccagag taaaaactaa tttacagaac ttccttattg atctgctggt tcttccagat 180

catattctgg ctattggtat ggctggcctt tctgaaggta ccctgcttgt ctattttcct 240
gactcagctc ttgcctgcct ttttcacatg ttgctgcaat tagactcacc gtgaggacta 300
cagtcaattt cagtctatct tgngcccaat acaacaagga tttttaatag tnncaaccca 360
cacctcaccc actaggactn aatgttcaca acangaagga ccattgctgc atactncttg 420
accancaact tttttgaaga tatttttaag tgcngagtag g 461

<210> 277
<211> 549
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(549)
<223> n = A,T,C or G

<400> 277
gggaagatgg cggacattca gactgagcgt gcctaccaaa agcagccgac catctttcaa 60
aacaagaaga gggtcctgct gggagaaact ggcaaggaga agctcccgcg gtactacaag 120
aacatcggtc tgggcttcaa gacacccaag gaggctattg agggcaccta cattgacaag 180
aaatgcccct tcactggtaa tgtgtccatt cgagggcgga tcctctctgg cgtggtgacc 240
aagatgaaga tgcagaggac cattgtcatc cgccgagact atctgcacta catccgcaag 300
tacaaccgct tcgagaagcg ccacaagaac atgtctgtac acctgtcccc ctgcttcagg 360
gacgtccaga tcggtgacat cgtcacagtg ggcgagtgcc ggcctctgag caagacagtg 420
cgcttcaacg tgctcaaggt caccaangct gccggcacca agaagcagtt ccagaagttc 480
tgangctgga catcggcccg ctccccacaa tgaaataaag ttattttctc attcccaaaa 540
aaaaaaaaa 549

<210> 278
<211> 344
<212> DNA
<213> Homo sapiens

<400> 278
ctgtagtccc agttactcgg gaggctgagg caggagaatc gcttgaaccc gggaggtgga 60
gattgcagtg agcccagatc gcaccactgc actccagtct ggcaacagag caagactcca 120
tctcaaaaag aaaagaaaag aagactctga cctgtactct tgaatacaag tttctgatac 180
cactgcactg tctgagaatt tccaaaactt taatgaacta actgacagct tcatgaaact 240
gtccaccaag atcaagcaga gaaaataatt aatttcatgg gactaaatga actaatgagg 300
ataatatttt cataattttt tatttgaaat tttgctgatt cttt 344

<210> 279
<211> 145
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(145)
<223> n = A,T,C or G

<400> 279
ccaacttggg gggctgngtc cacccagccc gnccgtcctg tgggctgcac agctcacctt 60
gttccctcct gccccggttc gagagccgag tctgtgggca ctctctgcct tcatgcacct 120
gtcctttcta acacgtcgcc ttcaa 145

<210> 280
<211> 410

<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(410)
<223> n = A,T,C or G

<400> 280
ccattactga ttttcaatta atttatgcat aaatgagacc caaactatca ctaattttca 60
gctatatgaa ttgatagcca cttgacatca gtgaaaggta cagtagggag tagatgaaat 120
tgtattttta atgaaaaggc tttgatggga gattcaagat ttttggtttt ttttttttt 180
gagacagggt cttgccctgt cacccaggct cgagtgcact ggagtgatca cagctcactg 240
gccgcaagtg atcctcctgc cttggcccct taagtgccag ggttacaggc atgagctacc 300
atgcctggca gaaattcaag atttggataa acttacttct ttgccaagcc tgttcttcaa 360
gttattcana actgggtgta taccttgtcc tcatatgtat cttgtccctg 410

<210> 281
<211> 377
<212> DNA
<213> Homo sapiens

<400> 281
ccattttcat cctgggtggt tagggcccct gtgggagcag atgggcactg tcaccaattt 60
ggctctgcca ggacagggca ggccctgcgg cctctggctg aatcacccac tgtccactcc 120
agagggtcca tcgagaatag ctgtccaagc aaggctgtac ctacgtacaa actaaagcta 180
ccgctcattc atctgctgtc caggaaagct taggagacat tcctgccttt ctacatggaa 240
aaaaaaatag tacaagtttt ggaattttct gtaattaaac aaggcatatt catgtactac 300
atatttcagc actaaggcgg ttgcttcact ttatatctat ataaaaaaag tggtaaaagt 360
cttttccttt tgtgcag 377

<210> 282
<211> 529
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(529)
<223> n = A,T,C or G

<400> 282
agacattact ggttatagaa ttaccacaac ccctacaaac ggccagcagg gaaattcttt 60
ggaagaagtg gtccatgctg atcagagctc ctgcactttt gataacctga gtcccggcct 120
ggagtacaat gtcagtgttt acactgtcaa ggatgacaag gaaagtgtcc ctatctctga 180
taccatcatc ccagaggtgc cccaactcac tgacctaagc tttgttgata taaccgattc 240
aagcatcggc ctgaggtgga ccccgctaaa ctcttccacc attattgggt accgcatcac 300
agtagttgcg gcaggagaag gtatccctat ttttgaagat tttgtggact cctcagtagg 360
atactacaca gtcacagggc tggagccggg cattgactat gatatcagcg ttatcactct 420
cattaatggc ggcgagagtg cccctactac actgacacaa caaacggntg ttcctnctnc 480
cactgacctg cgattcacca acattggtcc agacaccatg cgtgtcacc 529

<210> 283
<211> 558
<212> DNA
<213> Homo sapiens

<220>

<221> misc_feature
<222> (1)...(558)
<223> n = A,T,C or G

<400> 283
ccagcacctc tagaagggaa tttcttttct taatatcaca gttcccaaac ttaagacatt 60
atgatcaaga ttttctatct ttatttgcac tagtatctaa aatactcagt gaaatctttt 120
ttgctacact atatttactt ttacagactt tccattacca catacataca ccatcatgct 180
aagaaaccac caagtttttc ttctaatccc ccactaaaat taacaggttt caacaaactt 240
gaaattatag gggaactatg gggaaaacca gagaagtata tggaagaagg aagaagtgtg 300
aataggtcct acagaatttt acaatcactt tgccaagaca actataaata ctatgaataa 360
ttacttgaaa tcaggttgtg tagaatctat agttctctta aaaacaagtt ttgattctca 420
atattgcatt tttataccaa ataaaaagga tttagatcta acgtatttta gtngcatact 480
tactacctgc anactaaatt catttctcan gtactctaaa aaacttcaat agaacaaact 540
ttatgagatg ctataact 558

<210> 284
<211> 356
<212> DNA
<213> Homo sapiens

<400> 284
aaaaaataaa tggtatctta tttaattgtc ctgttccttc ccactccccg cctcctagga 60
tgttagccca agctcagggt aggcccaggg ggctgggaga aatgaagcca cccatgggga 120
ctggggacca ggggccttca gcatggcttc taggttccct cctcccccta ccccatctcc 180
tacctccaca gtacagactg tccccaactt aacagtggtt caacttaaac catgtttcaa 240
ctttacaatt ggtctgttgg ggtattaaat gaatttgtga cttaggatat tttcatttat 300
gatgggttta tcaggaagta accccatggt aagttgaggc atatctgtat atattt 356

<210> 285
<211> 184
<212> DNA
<213> Homo sapiens

<400> 285
ctggactagt agaaactcgc tgggaaggtg gtctgaagcc aggtgccttt gagttatcag 60
ggtgcatgtt ttccaagtgt ccaagcactg agttacccag gaacgctgac tgaacagtga 120
aagaggcatc tgtagcaact cgtgaggaca gtggaccatc tccccagccc tggttagctg 180
gcac 184

<210> 286
<211> 537
<212> DNA
<213> Homo sapiens

<400> 286
ctgttacagt gacaagagat aaaaagatag acctgcagaa aaaacaaact caaagaaatg 60
tgttcagatg taatgtaatt ggagtgaaaa actgtgggaa aagtggagtt cttcaggctc 120
ttcttggaag aaacttaatg aggcagaaga aaattcgtga agatcataaa tcctactatg 180
cgattaacac tgtttatgta tatggacaag agaaatactt gttgttgcat gatatctcag 240
aatcggaatt tctaactgaa gctgagatca tttgtgatgt tgtatgcctg gtatatgatg 300
tcagcaatcc caaatccttt gaatactgtg ccaggatttt taagcaacac tttatggaca 360
gcagaatacc ttgcttaatc gtagctgcaa agtcagacct gcatgaagtt aaacaagaat 420
acagtatttc acctactgat ttctgcagga aacacaaaat gcctccacca caagccttca 480
cttgcaatac tgctgatgcc cccagtaagg atatctttgt taaattgaca acaatgg 537

<210> 287
<211> 342

<212> DNA
<213> Homo sapiens

<400> 287
gtgctcgggg taatgacggt gctcgaggca gtgatggtca accaggccct cctggtcctc 60
ctggaactgc cggattccct ggatcccctg gtgctaaggg tgaagttgga cctgcagggt 120
ctcctggttc aaatggtgcc cctggacaaa gaggagaacc tggacctcag ggacacgctg 180
gtgctcaagg tcctcctggc cctcctggga ttaatggtag tcctggtggt aaaggcgaaa 240
tgggtcccgc tggcattcct ggagctcctg gactgatggg agcccggggt cctccaggac 300
cagccggtgc taatggtgct cctggactgc gaggtggtgc ag 342

<210> 288
<211> 562
<212> DNA
<213> Homo sapiens

<400> 288
aaatgcagtc cactctgctt tttgaagagg ctttggttca gctcccaaat ctcgattgct 60
tgacgcagtc tcctatgaga atactcagaa ggtgtcttct taaacaacaa acctattttt 120
agtggtggag ccgctcttag tagctgtgtc tgcgtgggac tgataaccaa tcactatctt 180
tggaggaagt cctaaccttt ccttgtatac cctccctata tgtgtaacag cttctctgtt 240
ttcacattca gtagtccata ttgctatctt atcaccttta gctctaacat taacaacagc 300
gccacataca tcatcactgt agtcatcaaa agattctcca ataaggcaca gaagtgtctc 360
tagccaaaag cgatcgaggt cacttcgtct ctgctgtttg ttcaatgtaa ttagccatcg 420
tcctcccccgt ttgtttttct catcttccca cataggctca ataccatcct taaaaagtga 480
gtagtcacag ccaggcatta aattactaga caactggata tggttgtaca gagcccaaaa 540
gtcttcaaca gtatcaaact tg 562

<210> 289
<211> 422
<212> DNA
<213> Homo sapiens

<400> 289
aaacaaaaag ttttgggtct gtctttggag tatttgtaac ttctaaattt tgaaatgact 60
gaattaggaa tttggatgct tattctttta gtctgtttgc ctaaaaacca atttacaatc 120
tgactgtctc ttgggagagg gaggtgcctt gcaaactttc acattaagaa tgtgcctgag 180
gctgctttac tctggaatag tctcagatct aaaatttcct ctatataagg tggcatatgt 240
taagttttgc ttcattggac cgtttagaat gctatgtaaa atgttgccat tctgttagat 300
tgctaactat atacccatct ctgatttggc tctccttaag tgataggatt tgttattcta 360
aaggtgataa acttgaaaat atcagaatct gagttttact tgaaattttg cagaataccc 420
ag 422

<210> 290
<211> 564
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(564)
<223> n = A,T,C or G

<400> 290
ctgtccaatg gcaacaggac cctcactcta ttcaatgtca caagaaatga cgcaagagcc 60
tatgtatgtg gaatccagaa ctcagtgagt gcaaaccgca gtgacccagt caccctggat 120
gtcctctatg ggccggacac ccccatcatt tccccccccag actcgtctta cctttcggga 180
gcgaacctca acctctcctg ccactcggcc tctaacccat ccccgcagta ttcttggcgt 240

atcaatggga taccgcagca acacacacaa gttctcttta tcgccaaaat cacgccaaat 300
aataacggga cctatgcctg ttttgtctct aacttggcta ctggccgcaa taattccata 360
gtcaagagca tcacagtctc tgcatctgga acttctcctg gtctctcaga agtgtaacat 420
tctgagtcaa cagcagacag agagctggag taaagaagtc agtgggttac ttgggagtga 480
tcagcctgac tctgaaatga cttttgatac caacataaag caagagtctg ggtcttctac 540
ttcttcatac agtggctatg aang 564

<210> 291
<211> 536
<212> DNA
<213> Homo sapiens

<400> 291
ttggacctcc tggctctctg ctgtacatcc gtggatccat catgtccatt ttgagacggg 60
aagatagtct tcaggaaaga cacacaaagg taacttgtgc agagggagat ggcaaattta 120
taacttctca gaaacacagt aatgataagt aaccaaggac ttccaccaaa gtcagtccca 180
cgatgacgat ggtcagccag agtattgata acctgatttc tggtcctccc caaccagctc 240
cctgtccctg cttctgggtg ctccttcctt cctgagctcc cagggttcct caaggtcact 300
tttggcgaca aaacataaaa aacaaatgat ggcaggatgg caggaagaac ctcataccca 360
agcagagtgc caggttttac agcctccgct cagccattca tatcctaagc aacaaaacat 420
cagcaggatg cggaaggtcc cgatagtaaa ccatctccat cacatccatg tagccatccg 480
tccatcaacc tcttagaatc atccagaaac aagtcactct tcatctgtcc agcaaa 536

<210> 292
<211> 578
<212> DNA
<213> Homo sapiens

<400> 292
ctgccatcac atcggacata ttggaggccc ttggaagaga cggtcacttc acactctttg 60
ctcccaccaa tgaggctttt gagaaacttc cacgaggtgt cctagaaagg atcatgggag 120
acaaagtggc ttccgaagct cctatgaagt accacatctt aaatactctc cagtgttctg 180
agtctattat gggaggagca gtctttgaga cgctggaagg aaatacaatt gagataggat 240
gtgacggtga cagtataaca gtaaatggaa tcaaaatggt gaacaaaaag gatattgtga 300
caaataatgg tgtgatccat ttgattgatc aggtcctaat tcctgattct gccaaacaag 360
ttattgagct ggctggaaaa cagcaaacca ccttcacgga tcttgtggcc caattaggct 420
tggcatctgc tctgaggcca gatggagaat acactttgct ggcacctgtg aataatgcat 480
tttctgatga tactctcagc atggatcagc gcctccttaa attaattctg cagaatcaca 540
tattgaaagt aaaagttggc cttaatgagc tttacaac 578

<210> 293
<211> 281
<212> DNA
<213> Homo sapiens

<400> 293
ctgagtgcca ggcgtggggt gacccccatt acgtcactct ggatgggcac cgattcgatt 60
tccaaggcac ctgcgagtac ctgctgagtg caccctgcca cagaccaccc ttgggggctg 120
agaacttcac tgtcactgta gccaatgagc accggggcag ccaggctgtc agctacaccc 180
gcagtgtcac cctgcaaatc tacaaccaca gcctgacact gagtgcccgc tggccccgga 240
agctacaggt cgacggcgtg ttcgtggctc tgcctttcca g 281

<210> 294
<211> 187
<212> DNA
<213> Homo sapiens

<400> 294

ctggtggcag gccaggccct cgcccacaca ctcgtcctct ggccggttgg cagtgtggag 60
cagagcttgg tgcgggttcc gaaagagctg gtcccagggc accgtgtgca cgaagcagag 120
gtgggtgtta tggtggatga gggccagtcc actgcccagt tccctcagtg agcgcagccc 180
cagccag 187

<210> 295
<211> 306
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(306)
<223> n = A,T,C or G

<400> 295
ctgcggtttg ntcaatggag ctgctgattg gggaaataat tttcaacact atcctgaatt 60
atgtgcctgt ctagataagc agagaccatg ccaaagctat aatggaaaac aagtttacaa 120
agagacctgt atttctttca taaaagactt cttggcaaaa aatttgatta tagttattgg 180
aatagcattt ggactggcag ttattgagat actgggtttg gtgttttcta tggtcctgta 240
ttgccagatc gggaacaaat gaatctgtgg atgcatcaag ctatcgtcag tcaaancccct 300
ttacct 306

<210> 296
<211> 381
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(381)
<223> n = A,T,C or G

<400> 296
gcgggatggg gatgatgagg ctattgtttt ttgtgaattc ttcgataatg gcccatttgg 60
gcaaaaagcc ggttagcggg ggcaggcctc ctagggagag gagggtggat ggaattaagg 120
gtgttagtca tgttagcttg tttcaggtgc gagatagtag tagggttgtg gtgctggagt 180
ttaagttgag tagtaggaat gcggtagtag ttaggataat ataaatagtt aaattaagaa 240
tggttatgtt agggttgtac ggtagaactg ctattattca tcctatgtgg gtaattgagg 300
agtatgctaa gattttgcgt anctgggttt ggtttaatcc acctcaactg cctgctatga 360
tggataagat tgagaacctc g 381

<210> 297
<211> 410
<212> DNA
<213> Homo sapiens

<400> 297
cgcttctgag ctaaaagctt ccatgaaggg gctgggaacc gacgaggact ctctcattga 60
gatcatctgc tccagaacca accaggagct gcaggaaatt aacagagtct acaaggaaat 120
gtacaagact gatctggaga aggacattat ttcggacaca tctggtgact tccgcaagct 180
gatggttgcc ctggcaaagg gtagaagagc agaggatggc tctgtcattg attatgaact 240
gattgaccaa gatgctcggg atctctatga cgctggagtg aagaggaaag gaactgatgt 300
tcccaagtgg atcagcatca tgaccgagcg gagcgtgccc cacctccaga aagtatttga 360
taggtacaag agttacagcc cttatgacat gttggaaagc atcaggaaag 410

<210> 298
<211> 260

<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(260)
<223> n = A,T,C or G

<400> 298
gctttttttt ttttttttt ttttttttt tgcacaatgg tttattaaag gaatgtatgg 60
cccacatcaa cctancaagg attctactgg taaaccttcc catggccaaa ggaaaaacaa 120
gcaggagttg agtggctggg gtggggtgca ggcaatggaa anagggcaaa agggtgtaaa 180
anctgaaggg ggctanaagc ttactcctga gtttnttcct tntgtcttna aatctttact 240
tnttatggcc aaanacccag 260

<210> 299
<211> 281
<212> DNA
<213> Homo sapiens

<400> 299
ccaaaaagat gctggggcag attgtggaca agtagaagca cctccttccc ctctgcgaca 60
ttgaatggcg tggattcaat agtgagcttg gcagtggtgg gcgggttcca gaaggttaga 120
agtgaggctg tgagcaggag cctctgccag gggatgcacc atctgtgggg aggggccgag 180
ggagactcca tggtctctgc tgtctgctct gtcctcctct gtggagaaga gcttgagttc 240
caggaacgtt ttgtcaaggc tgctgtgact gtctggtctg c 281

<210> 300
<211> 600
<212> DNA
<213> Homo sapiens

<400> 300
cctaccacaa taataaaaaa ccgtcaatta catcatcaca ttaaaataag ccagatgtac 60
aaaagtctga gacagtgaag acaaaaggac aacacaagat atttgttgaa aaatgtttgt 120
gctctttggg cacttaatta aacattgcaa aatcaacatc atcttcttct tcatcagact 180
ctgcaaaata ttttacttct ttcctagccc gaccggttcg tggcagagaa ggtggctcag 240
tagggaagtc tgaggggaag atgtccacat ctgaatcctg atcaaaagat gtcttcttcg 300
gtttcttgct tgttgttttg gatgttttcc tgccagggtt ataatcgcct tcattttcag 360
agccagatgc tttccttttc tttgcccctc ggcctttacc ttttggtgtt gtagtcttct 420
ttggaatgcc aaattctgaa tccgagtcag agtttacagc ctctactact ttcttctgtt 480
ttggggctct cttgggctta gggactgtat ctgaagacgg ttttcccttt ttagcagcta 540
ccgtttttac ttggaacttt atctgtctgt ttcagaccaa atgatggtga aaaaacagaa 600

<210> 301
<211> 305
<212> DNA
<213> Homo sapiens

<400> 301
ccttctctga aaaaagagaa ggaattactt attaaaacta agcacactta gcaacttctt 60
tcccaatcct atctttattc gtttgcctgg tgccaaattt ttctggccct ttttaatttg 120
caaaccttaa aaaaaaaaac aaagaaacaa aaacaccaaa cacacacata tctcacacat 180
agcactaagc tagaagcaga tataaatggg accactgtga atcaaagggg aaaaattcca 240
ggaaaaaaaa attccaatag cttcacagtt taactgaggt tttggaaaaa cttaagtgaa 300
ttcag 305

<210> 302

<211> 222
<212> DNA
<213> Homo sapiens

<400> 302
ccaaacttgc atttgcattt tgcactcatg acgatgatga tgcccatggc gcacagaacc 60
ccagcgcaga tgagcccgcc aacctggagg ctgtgccagt catagtagaa aggactgttt 120
ttatcttcta ggtcattggc gtccaggaca ggaaagcctg ccaggaacac aagcaggccc 180
agggtcacct tctgcatgtc agagcgctgg cctgtgtggt tc 222

<210> 303
<211> 195
<212> DNA
<213> Homo sapiens

<400> 303
ctgattttat ttccttctca aaaaaagtta tttacagaag gtatatatca acaatctgac 60
aggcagtgaa cttgacatga ttagctggca tgatttttc ttttttttcc cccaaacatt 120
gtttttgtgg ccttgaattt taagacaaat attctacacg gcatattgca caggatggat 180
ggcaaaaaaa agttt 195

<210> 304
<211> 172
<212> DNA
<213> Homo sapiens

<400> 304
ttgttttgtc tcttccttaa agcatttgca acagctacag tctaaaattg cttctttacc 60
aaggatattt acagaaaaga ctctgaccag agatcgagac catcctagcc aacatcgtga 120
aaccccatct ctactaaaaa tacaaaaatg agctgggctt ggtggcgcgc ac 172

<210> 305
<211> 146
<212> DNA
<213> Homo sapiens

<400> 305
ctgaaaaggg tggcagagaa ggaggcagtc aataagatgt ccctgcacaa cctcgccacg 60
gtctttggcc ccacgctgct ccggccctcc gagaaggaga gcaagctccc tgccaacccc 120
agccagccta tcaccatgac tgacag 146

<210> 306
<211> 377
<212> DNA
<213> Homo sapiens

<400> 306
ctgtttacag aaatatagtt gcgagtatac aaatgttcca atagaagcaa aatatctttt 60
taatatttaa caagttatca cagatagcta aaaacataga tgcaaatgaa attcccccag 120
agaacaaact gaaaatatct ggtatcagtg ctctgaaatc ccaactatga aagccatata 180
cacaaaaatg taacccttat atcattgcag gacaatggaa gaaggcagtt cagtggttga 240
tcagtgtgct caagcaaata aaattaaata aaaattaaaa atggcagaat ggtagctaaa 300
ccacttgaga acaggttaat gaaattattg gtactatact taaaacatta agtaaaagaa 360
gtgaatgaaa ctcattt 377

<210> 307
<211> 246
<212> DNA

<213> Homo sapiens

<400> 307
aaaacagtgt caaacttctc atgcatggag catgaattct tactaataga gactgtagtt 60
ttttttcttg tctttggtaa atatataaaa gaccttaatt tttctttttt aatgaatgga 120
gaaaacatga gaaaaccaga tggacctgtt agtactacat ttttaaggca ttttatattt 180
gatggtgccg tacttttaat aataataaaa ctgaagtttt ttagtggcaa tactgattta 240
tttttt 246

<210> 308
<211> 191
<212> DNA
<213> Homo sapiens

<400> 308
aaacgcaaag tagttggctg ggcagggctt gatgaggcca cacttgtact ttttaacctg 60
gatctccttg gtgggcgagg ctgccagcca gcgcggcaga cggatggtct tcatgctgaa 120
gctcatgtag cttcgaataa acatccatgt cgtgactatg gcaaagatga gggccaggag 180
gcgaagcaca c 191

<210> 309
<211> 342
<212> DNA
<213> Homo sapiens

<400> 309
ctgtgtgccc ctcctacatc aggggtaagg cccagctccc catcagcttc cttgaactgt 60
aaatcagatt agatttgggg atctgggctc agtctcagga gcagataaaa ctgggacact 120
cagccttggg gaagacaaag aaaagccaca taggaaagag atagacagac catgggcaag 180
ggaagattgc acagggaatg tgacatcagg gaacagatga gggaggagga ggcgcggcgg 240
cctcggggag aggacgggaa gcctgtcagg aaggggcccg ggaagcagga ggaggtgtgt 300
gtcatcgatg ccctgctggc tgacatcagg aagggcttcc ag 342

<210> 310
<211> 381
<212> DNA
<213> Homo sapiens

<400> 310
ccagtttgcc ctctcaggct cctgggatgg aaccctgcgc ctctgggatc tcacaacatt 60
gtctctggat ctcgagataa aaccatcaag ctatggaata ccctgggtgt gtgcaaatac 120
actgtccagg ctggagtgca gtagtgcgat ctcggctcac tgcaagctct gcttcccggg 180
ttcacgccat tctcctgcct cagcctcccg agtcgctggg actacaggcg cctgccatca 240
ggatgagagc cactcagagt gggtgtcttg tgtccgcttc tcgcccaaca gcagcaaccc 300
tatcatcgtc tcctgtggct gggacaagct ggtcaaggta tggaacctgg ctaactgcaa 360
gctgaagacc aaccacattg g 381

<210> 311
<211> 240
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(240)
<223> n = A,T,C or G

<400> 311

caaccgtggc atcncgcgaa tncggggcac cagctaccag agccctcacg gcatccccat 60
annacctgct ggaccggctg cttatcgnct ccaccacccc ctacagcgag aaagacacga 120
agcagatcct ccgcatccgg tgcgaggaag aagatgtgga gatgagtgag gacncctaca 180
cggtgctgac ccgcatcggg ctggagacgt cactgcgcta cgccatccag ctcatcacag 240

<210> 312
<211> 263
<212> DNA
<213> Homo sapiens

<400> 312
ctggagaagg agtgaatcct ccccatgtac ttgctacctg agaagcagtt gacataggat 60
gtacagatgg gtgaaaagtt ggctgctgca aatcaagcag atcattgggc agctttgatg 120
tgctgttaga agaactaggg gtagaaaata tgtcaatggc tggtgcagtc attatccctc 180
ctgctgaggt ggatacagga gaggctgcag ttgttaaaga ggtatgaggt ttctttgcaa 240
gttcttttag gcgctgttcc ttt 263

<210> 313
<211> 300
<212> DNA
<213> Homo sapiens

<400> 313
aaacaagatt tgctgcattt ccggcaatgc cctgtgcatg ccatggtccc tagacacctc 60
agttcattgt ggtccttgtg gcttctctct ctagcagcac ctcctgtccc ttgaccttaa 120
ctctgatggt tcttcacctc ctgccagcaa ccccaaaccc aagtgccttc agaggataaa 180
tatcaatgga acgcagagat gaacatctaa cccactagag gaaaccagtt tggtgatata 240
tgagacttta tgtggagtga aaattgggca tgccattaca ttgctttttc ttgtttgttt 300

<210> 314
<211> 123
<212> DNA
<213> Homo sapiens

<400> 314
ctgcagcccc cgtctcggcc cccaccagtg gctatcagga gtttgtacat gcggtggagc 60
agggtggcac ccaggtcagt gcggtggtgg gcttgggtcc cccaggagag gctggttaca 120
agg 123

<210> 315
<211> 371
<212> DNA
<213> Homo sapiens

<400> 315
ggaagggatg gtgacaggaa gaggcgtggt gtcacctgtg gatactgagg aaaggctggt 60
gacaggaaga ggggtggcct gacctgtgga tgcagaggaa gtgtcggtga caggaagagg 120
cgtggtgtca cctgtggata ctgaggaaag gctggtgaga ggaagagggg tggcgtgacc 180
ggtggatgct gaggaagcat cggtgacagg aagagtgctg gtgtcacctc tggatgctga 240
ggaagggctg gtgacatgaa gaggggtggc gtgacctgtg gataatgagg aagcattggt 300
gacaggaaga ggggtggtgt cacctgtgga tgctgaggaa gtgctggtga caggaacagg 360
ggtggtgtca c 371

<210> 316
<211> 270
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(270)
<223> n = A,T,C or G

<400> 316
ctggcctctg ccttcagggt accaccgtct ccaggacaca aatgggcagc agaaaaatgt 60
caccttgttg atactcagca gctcatctat tgggacaaaa cttccatctc ggccaaggga 120
aatactctgt tgagtgacca gcggggccca gcccccagcc ctatttatct catcaatatg 180
gttcanggaa gataaaaaag agtgttctat gggatagaaa ggtgggaata agaaaaaact 240
aagtggctgg gcacggtgag tcacgcctgt 270

<210> 317
<211> 344
<212> DNA
<213> Homo sapiens

<400> 317
ctgtagtccc agttactcgg gaggctgagg caggagaatc gcttgaaccc gggaggtgga 60
gattgcagtg agcccagatc gcaccactgc actccagtct ggcaacagag caagactcca 120
tctcaaaaag aaaagaaaag aagactctga cctgtactct tgaatacaag tttctgatac 180
cactgcactg tctgagaatt tccaaaactt taatgaacta actgacagct tcatgaaact 240
gtccaccaag atcaagcaga gaaaataatt aatttcatgg gactaaatga actaatgagg 300
ataatatttt cataattttt tatttgaaat tttgctgatt cttt 344

<210> 318
<211> 601
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(601)
<223> n = A,T,C or G

<400> 318
ctggaattgg cttacagcac atgctttgtt tcatgttatg ggtgaggacc tacatacact 60
cttactttag cagtcactta accttctcca gcaaggcagt tgtggggttc actaggattt 120
agtgcctgat ctttttttg ggaaggggcg ggaatgaatg tgttggggct gggagggaag 180
cagaagaaaa tgggagtgtg agtgagtgtg catgtgtctg aagttcacca ttgcccccac 240
ctgcacctag caaggaacag gtgtttgatg tattttgctc atgactgcag tatgcatgta 300
ttttttcct tctctgtgtt ttctaaactt acactaaagg attcatcaaa tcatcttgtt 360
cagatggctc aggattgtat ttattttgct taccccgtgc tcttgggttc tatagtattt 420
ctataattat gtaacgagaa tagtgttgca ctgtaatcta tcatatagag ctatatgtat 480
ggaaaatttt gancaatttt ttaagaaatg tatnctgttt gcaaaggcac agtaaagttt 540
gcatcttata gantataggc aaataaagct aanaattaaa ccttatttaa cacaaaccac 600
a 601

<210> 319
<211> 465
<212> DNA
<213> Homo sapiens

<400> 319
aaatgacttc agctaactta ccccaagaga aaaatctggc attgatctct tggttagcat 60
tattaccaat atcaactagc actattaaaa tcaaagttga aatggtacat tcatttgcca 120
aaaaaaaaaa aagaaaaaaa aggcttaaag gcaaagaagg tgaatcaacg tgcaaattag 180
catctggccc aattgcaaaa ttcatttcct ggatgtgagg gattgaacca tgcacacttg 240

caagcaagat gaagggcaaa cagatgacat caaatcaaaa ttaaccccac aaagaatcca 300
gaagacctca gatggtaaag gacagaggtc tacgtctact gctgttagtg ctcaggtatc 360
catcctgttt tccctaatct cctgattctg atccaagagt ctttgagacc tatgtcctag 420
gccatccttt catctagaaa tggaaaccat ggctgtggca ccagg 465

<210> 320
<211> 204
<212> DNA
<213> Homo sapiens

<400> 320
ccttgtgctg ggacaacctc tctcttgcct tacctcagag agggactatg ccctgacccc 60
tcctttctga aaatcagtgc cctccctgtt gctctaggag gctcctgctg gcttggtaga 120
agacagaatt cgatctgcct gtccctttt cccctggggt ttgacacaca ggctcctctc 180
agcatgaggt ggagcagtga ccag 204

<210> 321
<211> 420
<212> DNA
<213> Homo sapiens

<400> 321
ctcgcgtcgc atttggccgc ctccctaccg ctccaagccc agccctcagc catggcatgc 60
cccctggatc aggccattgg cctcctcgtg gccatcttcc acaagtactc cggcagggag 120
ggtgacaagc acaccctgag caagaaggag ctgaaggagc tgatccagaa ggagctcacc 180
attggctcga agctgcagga tgctgaaatt gcaaggctga tggaagactt ggaccggaac 240
aaggaccagg aggtgaactt ccaggagtat gtcaccttcc tgggggcctt ggctttgatc 300
tacaatgaag ccctcaaggg ctgaaaataa atagggaaga tggagacacc ctctgggggt 360
cctctctgag tcaaatccag tggtgggtaa ttgtacaata aatttttttt ggtcaaattt 420

<210> 322
<211> 314
<212> DNA
<213> Homo sapiens

<400> 322
ctccgcccag ccagatgtcc cgagtgcgcc aaggactgtc ctctcaccca ctcctggatt 60
ctgccctgac ctccatcctg gacactgcct taataacata gacccttcca ctgacaccct 120
cgctctcaca ccccctccag ggcaggggcc cttagagtct tggttgccaa acagatttgc 180
agatcaagga gaacccagga gtttcaaaga agcgctagta aggtctctga gatccttgca 240
ctagctacat cctcagggta ggaggaagat ggcttccaga agcatgcggc tgctcctatt 300
gctgagctgc ctgg 314

<210> 323
<211> 423
<212> DNA
<213> Homo sapiens

<400> 323
ccaccagctc atcgcaatca ttgacgatag gcagcttccc tttcttgcta cgctgcagga 60
tctcatttgc ctctttcaac gtcacacctg ctggagccac caccagttca atccttggcg 120
tcatcacctc actgaggagg gtggtgtggt ccttctcagc aagaaagtcg atgtctcggg 180
aggtgacgat gcccaccagc ttgctgccca tggtgcccgt ctcagtgatg gggatgccag 240
agaagccatg ccgcatcttg gcctccagca catcgcccac agtgtgcgag gggctcagca 300
ccacagggtc cgtgatgaag ccctgttcaa acttcttgac cttccgcacc tcgttggcct 360
ggaactctgg ggtgcagttg tggtgaatga aaccaatacc tcccatcaga gccatggcaa 420
tgg 423

<210> 324
<211> 427
<212> DNA
<213> Homo sapiens

<400> 324
ctgcatcgcg gcccacgaca agaggggggag gtacgggacc ctgttcacga tggaccgggt 60
gctgaccccc caatggggac tgtcatggat gtcctgaagg gagacaatcg ctttagcatg 120
ctggtagctg ccatccagtc tgcaggactg acggagaccc tcaaccggga aggagtctac 180
acagtctttg ctcccacaaa tgaagccttc cgagccctgc caccaagaga acggagcaga 240
ctcttgggag atgccaagga acttgccaac atcctgaaat accacattgg tgatgaaatc 300
ctggttagcg gaggcatcgg ggccctggtg cggctaaagt ctctccaagg tgacaagctg 360
gaagtcagct tgaaaaacaa tgtggtgagt gtcaacaagg agcctgttgc cgagcctgac 420
atcatgg 427

<210> 325
<211> 401
<212> DNA
<213> Homo sapiens

<400> 325
ctggtaaccc ttccacaccc caatcttcat ggaccagaga tcttggatgt tccttccaca 60
gttcaaaaga ccccttttcgt cacccaccct gggtatgaca ctggaaatgg tattcagctt 120
cctggcactt ctggtcagca acccagtgtt gggcaacaaa tgatctttga ggaacatggt 180
tttaggcgga ccacaccgcc cacaacggcc acccccataa ggcataggcc aagaccatac 240
ccgccgaatg taggacaaga agctctctct cagacaacca tctcatgggc cccattccag 300
gacacttctg agtacatcat tccatgtcat cctgttggca ctgatgaaga acccttacag 360
ttcagggttc ctggaacttc taccagtgcc actctgacag g 401

<210> 326
<211> 263
<212> DNA
<213> Homo sapiens

<400> 326
ctggagaagg agtgaatcct ccccatgtac ttgctacctg agaagcagtt gacataggat 60
gtacagatgg gtgaaaagtt ggctgctgca aatcaagcag atcattgggc agctttgatg 120
tgctgttaga agaactaggg gtagaaaata tgtcaatggc tggtgcagtc attatccctc 180
ctgctgaggt ggatacagga gaggctgcag ttgttaaaga ggtatgaggt ttctttgcaa 240
gttcttttag gcgctgttcc ttt 263

<210> 327
<211> 344
<212> DNA
<213> Homo sapiens

<400> 327
ctgtccaatg acaacaggac cctcactcta ctcagtgtca caaggaatga tgtaggaccc 60
tatgagtgtg gaatccagaa caaattaagt gttgaccaca gcgacccagt catcctgaat 120
gtcctctatg gcccagacga ccccaccatt tccccctcat acacctatta ccgtccaggg 180
gtgaacctca gcctctcctg ccatgcagcc tctaacccac ctgcacagta ttcttggctg 240
attgatggga acatccagca acacacacaa gagctcttta tctccaacat cactgagaag 300
aacagcggac tctatacctg ccaggccaat aactcagcca gtgg 344

<210> 328
<211> 512
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(512)
<223> n = A,T,C or G

<400> 328
gaatgcggct gttaagacct gcaataatcc agaatggcta ctctgatcta tgttgataag 60
gaaaatggag aaccaggcac ccgtgtggtt gctaaggatg ggctgaagct ggggtctgga 120
ccttcaatca aagccttaga tgggagatct caagtttcaa caccacgttt tggcaaaacg 180
ttcgatgccc caccagcctt acctaaagct actagaaagg ctttgggaac tgtcaacaga 240
gctacagaaa agtctgtaaa gaccaaggga cccctcaaac aaaaacagcc aagcttttct 300
gccaaaaaga tgactgagaa gactgttaaa gcaaaaagct ctgttcctgc ctcagatgat 360
gcctatccag aaatagaaaa attctttccc ttcaatcctc tagactttga gagttttgac 420
ctgcctgaag agcaccagat tgcgcacctc cccttgagtg gagtgcctct catgatcctt 480
gacnangaga gagancttga aaagctgttt ca 512

<210> 329
<211> 364
<212> DNA
<213> Homo sapiens

<400> 329
ctgtgttcct aaagttcgtc tttcgcttgg ctcaggacaa agcggtgtaa cgagtcaagg 60
tctctgcctc cactgtgctc actgactttc ttccctcctc ggaaaagcaa taacgtgggg 120
tagcctcgta ccgaatactt gctgcagata ttccgttcag cagtgcagtc tacttcggcg 180
atcttgaccc ccgccagacc agggaattcc tttttagaga gttcctccca agtaggagcc 240
agagtcttac aatgaccaca ccatggagca taaaacttga tgaaggttat tccttctgca 300
atggtgtcat cgaagttatt ttcagtgagt gccaacacag tgcccttgtc agcctcgggc 360
tcag 364

<210> 330
<211> 221
<212> DNA
<213> Homo sapiens

<400> 330
cacccggccg ccttctaact gtgactcccc gcactcccca aaaagaatcc gaaaaaccac 60
aaagaaacac caggcgtacc tggtgcgcga gagcgtatcc ccaactggga cttccgaggc 120
aacttgaact cagaacacta cagcggagac gccacccggt gcttgaggcg ggaccgaggc 180
gcacagagac cgaggcgcat agagaccgag gcacagccca g 221

<210> 331
<211> 520
<212> DNA
<213> Homo sapiens

<400> 331
gttggtgcac aaaatactgc catttgctca aagctggctg ccaaatgttt ggtgatgaag 60
gcagaaatga atggctcaaa acttgggaga agagcaaaac ctgaaggggc cctccagaac 120
aatgatgggc tttatgatcc tgactgcgat gagagcgggc tctttaaggc caagcagtgc 180
aacggcacct ccacgtgctg gtgtgtgaac actgctgggg tcagaagaac agacaaggac 240
actgaaataa cctgctctga gcgagtgaga acctactgga tcatcattga actaaaacac 300
aaagcaagag aaaaacctta tgatagtaaa agtttgcgga ctgcacttca gaaggagatc 360
acaacgcgtt atcaactgga tccaaaattt atcacgagta ttttgtatga gaataatgtt 420
atcactattg atctggttca aaattcttct caaaaaactc agaatgatgt ggacatagct 480
gatgtggctt attattttga aaaagatgtt aaaggtgaat 520

<210> 332
<211> 305
<212> DNA
<213> Homo sapiens

<400> 332
ccaccccgga gatgacacga ggctcacatg actctagaca cttggtggaa agtgaggcga 60
gaaaaacaat gacttgggcc aattacacga ctgcaaagct agagctgcca acagggctcc 120
agggagcttg gcttctgtag aagttctaag gaagcggtac gaactccacg gcggtggggc 180
gctaactagc agggacccct gcaagtgttg gtcgggggcc tcgagctgcc tgagctgaca 240
cgaggggggag gggtctgtgt agccaacagg tgaccgaagg gcttgcctgc ccacagctta 300
cttgg 305

<210> 333
<211> 377
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(377)
<223> n = A,T,C or G

<400> 333
ccttaaatat cagactttgt aacaggtaga aatattacag aataatttaa gacactacaa 60
tgggggcaaa tgaaatnnga aaatttttan tgagttaccc gtactcatta cattttcagt 120
gcttttacaa ggaaaaaagg tgatatgttt aattttaaca ttttaattgg ctagctcttg 180
cccttatatg actttaatgt ctgtgagtca ttcccagctt aaattaacaa ttgttagtat 240
tagtctcaca cataagtgcc atacatttta tcctcatgga tgtgatgcac tgaaaagtta 300
gttgctctcc ttttttcttt tttttgtcgt gcatatttta tttctgtagt ttctggntag 360
ctaccctaaa gtgattt 377

<210> 334
<211> 251
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(251)
<223> n=A,T,C or G

<400> 334
cagtcaaaaa cttcacncaa tnggaaaata aangnttntt caatgaataa tcaaacaaaa 60
attatccagg accttatagg gttttcagta tgtnccaggc ttgatgcnca tnttaaaana 120
caggacatta tnttgctggg atcattaggg tatgactgaa agngaaaaac agtaatttgt 180
aaaacattta cctaataata gctttcccaa acagtacttc ccctggaatt aaaacaggaa 240
atacaattta t 251

<210> 335
<211> 513
<212> DNA
<213> Homo sapiens

<400> 335
aaaaagaaaa aaaaagccaa atacattttc tgacattgta agattgcctt actgtctgtc 60

attccttatt gctggcccct ttctcaggcc ggaggccaag tggtggagaa ggaaaggaaa 120
tgatcgaacg ggcatgttgt caagtgggca tgccactggg aaataccacc agtttacact 180
gaaacattgt cctcagagga gtaggaaagt ggattttgaa tctctatttt gctcaaaagt 240
tcagttcctg agatactgat gactgagagt gctgctggga aattttcagg attgtgtggt 300
cttttggggt tttttgtttt tttttttaa gacaaagttg accgctgttc actgtccacg 360
tgatcagttg taagattaca atgctgcatg ctagttggtt acataagata caattccagt 420
gatggaaggc ggttataatg gatggtggtg tgtacaagat ggcactgcca tctttgagca 480
gagcccagct ctgcagcgcc acttcatctt ttt 513

<210> 336
<211> 343
<212> DNA
<213> Homo sapiens

<400> 336
aaatttttta ggttaatttt cttgctgtga tatatatgag gaatttacta ctttatgtcc 60
tgctctctaa actacatcct gaactcgacg tcctgaggta taatacaaca gagcactttt 120
tgaggcaatt gaaaaaccaa cctacactct tcggtgctta gagagatctg ctgtctccca 180
aataagcttt tgtatctgcc agtgaattta ctgtactcca aatgattgct ttcttttctg 240
gtgatatctg tgcttctcat aattactgaa agctgcaata ttttagtaat accttcggga 300
tcactgtccc ccatcttccg tgttagagca aagtgaagag ttt 343

<210> 337
<211> 647
<212> DNA
<213> Homo sapiens

<400> 337
cctctgtgca agcarcacat aggatctgga tgtaggttga ggataratcc tcacccacca 60
gkggggtaay tttcccagca attctgaaac taaaataagg aaggcacwtt cccarasccc 120
tgytgagtag gggcttcagg ctattttcac tctacacaaa atgggggaga ggagtyccct 180
ctccactaat ttttcaccca taaacctcca catcactagg aacctaaagg ggaactccaa 240
aggccaacac atccttggtg gttatatgtg ttgtcctgac aacctcctgc tccagaaatg 300
ccaggagcat tggatatgtc attgggagca tcaggcagtc caacatcgga gggagaaagg 360
cccagagatg aggatctgag tcaggctggc aaggctggag tcagaaagtg accattaggc 420
aactggtcac tacaattggt ggctacaaag aagtggtcac agtcaccaaa ataaagaggt 480
ttacaacaac ggtttcccct taggtcattt tgaccaggac agtaccctaa aggaaataag 540
gcagcatcgc ataaagcaag agccccctc agtaatccac ggagatggag gggtcctcat 600
cctggttgca aaacaggaga tggaaaggcc caggggtggc gactcag 647

<210> 338
<211> 515
<212> DNA
<213> Homo sapiens

<400> 338
aaaaaaargy wwkktctyca ctccaaaatg acagagacag actactgaag gaattgaaga 60
atctgcagca gcaacmcyta cagattawyc arrrgwycac tgagttacat ccactgaagg 120
ctcaacttca ggagtatcaa gataagacra aagcatttca gattatgcaa gaagagctca 180
ggcaggaaaa cctctcctgg cagcatgagc tgcatcagct caggatggag aagagttcct 240
gggaaataca tgagaggaga atgaaggaac agtaccttat ggctatctca gataaagatc 300
agcagctcag tcatctgcag aatcttataa gggaattgag gtcttcttcc tcccagactc 360
agcctctcaa agtgcaatac caaagacagg catccccaga gacatcagct tccccagatg 420
ggtcacaaaa tctggtttat gagacagaac ttctcaggac ccagctcaat gacagcttaa 480
aggaaattca ccaaaaggag ttaagaattc agcaa 515

<210> 339
<211> 438

<212> DNA
<213> Homo sapiens

<400> 339
aaaataggcc ctgagtataa gagcatgaag agctgccttt atgtcggcat ggcgagcgac 60
aacgtcgatg ctgctgagct cgcggagacc attgcggcca cagcccggga gatagaggag 120
aactcgaggc ttctggaaaa catgacagaa gtggttcgga aaggcattca ggaagctcaa 180
gtggagctgc agaaggcaag tgaagaacgg cttctggaag aggggtgtt gcggcagatc 240
cctgtagtgg gctccgtgct gaattggttt tctccggtcc aggctttaca gaagggaaga 300
acttttaact tgacagcagg ctctctggag tccacagaac ccatatatgt ctacaaagca 360
caaggtgcag gagtcacgct gcctccaacg ccctcgggca gtcgcaccaa gcagaggctt 420
ccaggccaga agcctttt 438

<210> 340
<211> 451
<212> DNA
<213> Homo sapiens

<400> 340
ctgatgatgt agaagtatat gattgaacga ccagagccag aattccaaga cctaaacgaa 60
aaggcacgag cacttaaaca aattctcagt aagatcccag atgagatcaa tgacagagtg 120
aggtttctgc agacaatcaa ggatatagct agtgcaataa aagaacttct tgatacagtg 180
aataatgtct tcaagaaata tcaataccag aaccgcaggg cacttgaaca ccaaaagaaa 240
gaatttgtaa agtactccaa aagtttcagt gatactctga aaacgtattt taaagatggc 300
aaggcaataa atgtgttcgt aagtgccaac cgactaattc atcaaaccaa cttaatactt 360
cagaccttca aaactgtggc ctgaaagttg tatatgttaa agagatgtac ttctcagtgg 420
cagtattgaa ctgcctttat ctgtaaattt t 451

<210> 341
<211> 237
<212> DNA
<213> Homo sapiens

<400> 341
aaaaccatca taacaaaaag ggtccattgt cttatgatcc actggaaaga ggaccgactc 60
atcatttatg gctatgactt ggcagtgact ccaatgtgat atcctgtaat tttatcttca 120
gttatgctat agcatgtaca tttccattct cttgtcgaag tttctttcgt tcctcagctt 180
ctccttcata tttcctgacg tattgtcttc taagctggac tgtaataaca gcaacag 237

<210> 342
<211> 512
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(512)
<223> n = A,T,C or G

<400> 342
tgtaaaacga cggccagtga attgtaatac gactcactat agggcgaatt gggccctcta 60
gatgcatgct ngagcggccg cccagtgtga tggatatctg cagaattcgc ccttgagcgg 120
ccgcccgggc gggtcctggg agatcccagg gtcctccacc ctccccctga ccacatacaa 180
aggcactcta gttcaagggt gaaaagtctc acccaggagg aacagccctc cttgaagcaa 240
tggcagggcc agcagggagg tgggcatggc agggaatgga gtgagccaga cagacttcac 300
ctccttactg gacacagggt caagggcgag tttcaattgc tgctcccttt actttctcta 360
cctgtgacta ctccctggac caatcctgag gagggcacat tttccagaag ccacgtgata 420
ggggctggtt tctgtggagc cggaggcaga gacactgaac ttgagctcac ctcctaacac 480

cggcagtaaa cttcctggaa ctttgccctc ag 512

<210> 343
<211> 372
<212> DNA
<213> Homo sapiens

<400> 343
aaatgttctc atcagtttct tgccatgttg ttaactatac aacctggcta aagatgaata 60
tttttctact ggtattttaa tttttgacct aaatgtttaa gcattcggaa tgagaaaact 120
atacagattt gagaaatgat gctaaattta tagttttcag taacttaaaa agctaacatg 180
agagcatgcc aaaatttgct aagtcttaca aagatcaagg gctgtccgca acagggaaga 240
acagttttga aaatttatga actatcttat ttttaggtag gttttgaaag ctttttgtct 300
aagtgaattc ttatgccttg gtcagagtaa taactgaagg agttgcttat cttggctttc 360
gagtctgagt tt 372

<210> 344
<211> 308
<212> DNA
<213> Homo sapiens

<400> 344
ggccagccca ccctcctggg gctgacatga gccattccct gtgatgttca ctctcctccc 60
aaagcaaacc acagccaagc ctgtctgagc tgggagtccc cttccccagc agagctccca 120
gtccctgcat acccagcggg gtggcgactc gggaagagct gagctggaga cggctctaga 180
ccaagtccgg ccaaccaggc ggttctgtaa tcctctccca gggcccatgg aagttaggct 240
tccatcaggc gcacttttcc caccaggggc tctgggagga cgtgtcttct aaagtgttcc 300
gtgctcac 308

<210> 345
<211> 513
<212> DNA
<213> Homo sapiens

<400> 345
gaatgcggct gttaagacct gcaataatcc agaatggcta ctctgatcta tgttgataag 60
gaaaatggag aaccaggcac ccgtgtggtt gctaaggatg ggctgaagct ggggtctgga 120
ccttcaatca aagccttaga tgggagatct caagtttcaa caccacgttt tggcaaaacg 180
ttcgatgccc caccagcctt acctaaagct actagaaagg ctttgggaac tgtcaacaga 240
gctacagaaa agtctgtaaa gaccaaggga cccctcaaac aaaaacagcc aagcttttct 300
gccaaaaaga tgactgagaa gactgttaaa gcaaaaagct ctgttcctgc ctcagatgat 360
gcctatccag aaatagaaaa attctttccc ttcaatcctc tagactttga gagttttgac 420
ctgcctgaag agcaccagat tgcgcacctc cccttgagtg gagtgcctct catgatcctt 480
gacgaggaga gagagcttga aaagctgttt cag 513

<210> 346
<211> 744
<212> DNA
<213> Homo sapiens

<400> 346
aaaaaataca ggagtcgata gcagcagttg gtgacgagat ggcactcaga aacggcgttg 60
acgtaattta ggacgtggaa tcataagcga aacagcacac tgtttgaata aagagcagag 120
tcggtattta tatttgktty tcttttgtca tgattatttg atttttaagk tgctccagct 180
aaggcatttt tttgtattag tatttctatt agggaacctt tcttattagg tggtttgtat 240
tgtctggttt ctaacatgca ggtagctgtt tggcagttaa acacgtttag agtaatttga 300
gttacaacgt gtgaaactga gcaaaaaagc agtgataagt ttgggttacc ataccaaata 360
tttgttttcc cactggaaaa aagtaagttt tagaaaatag ttaacctttg cagcatttgt 420

ttacagttta cagttccaga agtgcgtcga aatggattac ataactgctc ttttattcct 480
ggtgttcaca tctgtcccag gctgacacct gctcttggct ggcccacttt ggtatgggct 540
ttaatttcac taccccaaac acgatactgt catctgcttt ataataatgc tcaagatgcc 600
tgataaaaat ctcattttgc agccagacaa gccttgaatc cttttggcac taactgcaaa 660
ggaagatttt tttctctaga tatgcattag cagctagtgc tccagttaga agcacgaact 720
ataaccttga taagtaaaca gcag 744

<210> 347
<211> 392
<212> DNA
<213> Homo sapiens

<400> 347
ctggtgctag gttacgaggg ctggctggcc ggctaccaga tgaattttga gactgcaaaa 60
tcccgagtga cccagagcaa ctttgcagtt ggctacaaga ctgatgaatt ccagcttcac 120
actaatgtga atgacgggac agagtttggc ggctccattt accagaaagt gaacaagaag 180
ttggagaccg ctgtcaatct tgcctggaca gcaggaaaca gtaacacgcg cttcggaata 240
gcagccaagt atcagattga ccctgacgcc tgcttctcgg ctaaagtgaa caactccagc 300
ctgataggtt taggatacac tcagactcta aagccaggta ttaaactgac actgtcagct 360
cttctggatg gcaagaacgt caatgctggt gg 392

<210> 348
<211> 476
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(476)
<223> n = A,T,C or G

<400> 348
gattacgcca agcttgggta cccgagctcg gatccactag taacggccgc ccccgggctg 60
cargaattcm kaggtgccag aagccgagtt cctasacccc akctggattt ttttttttcc 120
aatttsgtgg cggcgggcgg ctggggaagc agctctgtgg tctcccggcg ggtgaggccg 180
agacagagac caacctctag gcggcgccct ggcgttgagt ccgcgctgaa caagattctc 240
tcgcagctct gcgtcccgcg ccgggtgaag gggcctgttc ctgcaggcgc tcggggctcg 300
gcctggagct ggcggccggc acgtcgcctt tggcgccctc tctcggaacc acgccgtccc 360
ccgcgcgtct tcgtagactg caggagtcca gggcgtctgg ggactgtgac ggcscccaag 420
gcgggggatg tggcggctnt gggtcgccgc cttctgcccc cgcctctgtg aggcct 476

<210> 349
<211> 732
<212> DNA
<213> Homo sapiens

<400> 349
cctgggtggt ggagcgaatg ggccgattcc accggatcct ggagcctggt ttgaacatcc 60
tcatccctgt gttagaccgg atccgatatg tgcagagtct caaggaaatt gtcatcaacg 120
tgcctgagca gtcggctgtg actctcgaca atgtaactct gcaaatcgat ggagtccttt 180
acctgcgcat catggaccct tacaaggcaa gctacggtgt ggaggaccct gagtatgccg 240
tcacccagct agctcaaaca accatgagat cagagctcgg caaactctct ctggacaaag 300
tcttccggga acgggagtcc ctgaatgcca gcattgtgga tgccatcaac caagctgctg 360
actgctgggg tatccgctgc ctccgttatg agatcaagga tatccatgtg ccaccccggg 420
tgaaagagtc tatgcagatg caggtggagg cagagcggcg gaaacgggcc acagttctag 480
agtctgaggg gacccgagag tcggccatca atgtggcaga agggaagaaa caggcccaga 540
tcctggcctc cgaagcagaa aaggctgaac agataaatca ggcagcagga gaggccagtg 600
cagttctggc gaaggccaag gctaaagctg aagctattcg aatcctggct gcagctctga 660

cacaacataa tggagatgca gcagattcac tgactgtggc cgagcagtat gtcagcgcgt 720
tctccaaact gg 732

<210> 350
<211> 938
<212> DNA
<213> Homo sapiens

<400> 350
ttttcaagct gacttccagc ttgtcacctt ggagagactt tagccgcacc agggccccga 60
tgcctccgct aaccaggatt tcatcaccaa tgtggtattt caggatgttg gcaagttcct 120
tggcatctcc caagagtctg ctccgttctc ttggtggcag ggctcggaag gcttcatttg 180
tgggagcaaa gactgtgtag actccttccc ggttgagggt ctccgtcagt cctgcagact 240
ggatggcagc taccagcatg ctaaagcgat tgtctccctt caggacatcc atgacagtcc 300
ccattggggg ggtcagcacc cggtccatcg tgaacagggt cccgtacctc cccctcttgt 360
cgtgggccgc gatgcagctg ttctcaatgc agaggctatt acgataaaca aaaactctca 420
gtttttgcc gcccagagtt tccagggtct gtccatggta cagatactta gaggccagct 480
ggtctttaat tatgtggttc cgaagcaaat tccttgtatg ggcatcaatt ggaggggttc 540
catctttgaa tacagaattc aggggagcca ggagggtcaa ccgctcactt ccagagagat 600
gattgccgag gccggcttgt ctgaaaaggt caatggctgt ggacacatca gactctgcag 660
ccaattcaaa tagtgtcttg gctgagtctg ggatgagtag ctcatcaatg tagtggatca 720
ccccgttggt ggctaggatg tctttattgg agatgatcgc cttcccgttg atagtgagca 780
tgtccccgct gcagcccacc tccagtgtcg tgccctccag ggtctctaca gacagccccg 840
caacgatggc ttcagcacac atagctaact tcaagatgtg gttgttcagc aggtctctca 900
gggcttctgg gtcgcccagg atacggttca aagtctca 938

<210> 351
<211> 793
<212> DNA
<213> Homo sapiens

<400> 351
aaaaatctac ctgttcctga cttaaaacaa aaggaaagaa actacctttt tataatgcac 60
aactgttgat ggtaggctgt atagttttta gtctgtgtag ttaatttaat ttgcagtttg 120
tgcggcagat tgctctgcca agatacttga acactgtgtt ttattgtggt aattatgttt 180
tgtgattcaa acttctgtgt actgggtgat gcacccattg tgattgtgga agatagaatt 240
caatttgaac tcaggttgtt tatgagggga aaaaaacagt tgcatagagt atagctctgt 300
agtggaatat gtcttctgta taactaggct gttaacctat gattgtaaag tagctgtaag 360
aatttcccag tgaaataaaa aaaaaatttt aagtgttctc ggggatgcat agattcatca 420
ttttctccac cttaaaaatg cgggcattta agtctgtcca ttatctatat agtcctgtct 480
tgtctattgt atatataatc tatatgatta aagaaaatat gcataatcag acaagcttga 540
atattgtttt tgcaccagac gaacagtgag gaaattcgga gctatacata tgtgcagaag 600
gttactacct agggtttatg cttaatttta atcggaggaa atgaatgctg attgtaacgg 660
agttaatttt attgataata aattatacac tatgaaaccg ccattgggct actgtagatt 720
tgtatccttg atgaatctgg ggtttccatc agactgaact tacactgtat attttgcaat 780
agttacctca agg 793

<210> 352
<211> 671
<212> DNA
<213> Homo sapiens

<400> 352
ccagttagat tatgggtcca aagggattcc agacacttct gagccagtca gctaccacaa 60
ctctggagta aaatatgctg catccgggca agaatcttta agactgaacc acaaagaggt 120
aaggctctcc aaagagatgg agcgaccctg ggttaggcag ccttctgccc cagagaaaca 180
ctccagagac tgctacaagg aggaagaaca cctcactcag tcaatcgtcc cacccctaa 240
accagagagg agtcatagcc tcaaactcca tcatacccag aacgtggaga gggaccccag 300

tgtgctgtac cagtaccaac cacacggcaa gcgccagagc agtgtgactg ttgtgtccca 360
gtatgataac ctggaagatt accactccct gcctcagcac cagcgaggag tctttggagg 420
gggcggcatg gggacgtatg tgccccctgg ctttccccat ccacagagca ggacctatgc 480
tacagcgttg ggtcaagggg ccttcctgcc cgcagagttg tccttgcagc atcctgaaac 540
acagatccat gcagaatgag ccctgcgagc aatagagttg aagcagcctc tgctggacag 600
tggactgttc tatttttttc aataaccaaa aagattaaac aaaaaatact ataaaacccc 660
tgaccacatt t 671

<210> 353
<211> 571
<212> DNA
<213> Homo sapiens

<400> 353
ccaacatggt gaaaccctct ctccactaaa aatacaaaaa ttagccagga atggtggcgg 60
gcgcctgtag tcccagctac ttgggaggct gaggcaggag aatcgcttga acccgggagg 120
tggaggctgc agtgagccag gatcgcgcca ctgcactcca gcctgggcaa caagagcgaa 180
actccatctc aaaaaaaaaa aaaaaagtgc tgttaataca atcaggatgg acaggaagta 240
aagaggtaaa aattctacgc accttgagat ataagccaga tttttaagac gaaaaccaag 300
attttgttta agaaaaaaag aaggaaggat tacgctgcag acgggccatt ctaggggtca 360
gcttccctcc cctccttccc tcttatcagc cagagacaga aactaaaaac cagggtttag 420
ggcagatgaa agcctaaaca gaaagaagga tgggggtcgg gagaaaggaa aaaacagcaa 480
ctgcctagat acaagcagag aagacaaagg cctcattcca ggtcagctgg gcaattctct 540
caggcttgcc atcttgtgtc ctgttttctt c 571

<210> 354
<211> 368
<212> DNA
<213> Homo sapiens

<400> 354
cctccaccat ggccttcaag cagatggagc agatctctca gttcctgcaa gcagctgagc 60
gctatggcat taacaccact gacatcttcc aaactgtgga cctctgggaa ggaaagaaca 120
tggcctgtgt gcagcggacg ctgatgaatc tgggtgggct ggcagtagcc cgagatgatg 180
ggctcttctc tggggatccc aactggttcc ctaagaaatc caaggagaat cctcggaact 240
tctcggataa ccagctgcaa gagggcaaga acgtgatcgg gttacagatg ggcaccaacc 300
gcggggcgtc tcaggcaggc atgactggct acgggatgcc acgccagatc ctctgatccc 360
accccagg 368

<210> 355
<211> 509
<212> DNA
<213> Homo sapiens

<400> 355
aaaattaggc tgaatgtact tcatgtgatt tgtcaaccat agtttatcag agattatgga 60
cttaattgat tggtatatta gtgacatcaa cttgacacaa gattagacaa aaaattcctt 120
acaaaaatac tgtgtaacta tttctcaaac ttgtgggatt tttcaaaagc tcagtatatg 180
aatcatcata ctgtttgaaa ttgctaatga cagagtaagt aacactaata ttggtcattg 240
atcttcgttc atgaattagt ctacagaaaa aaaatgttct gtaaaattag tctgttgaaa 300
atgttttcca aacaatgtta ctttgaaaat tgagtttatg tttgacctaa atgggctaaa 360
attacattag ataaactaaa attctgtccg tgtaactata aattttgtga atgcattttc 420
ctggtgtttg aaaaagaagg gggggagaat tccaggtgcc ttaatataaa gtttgaagct 480
tcatccacca aagttaaata gagctattt 509

<210> 356
<211> 241
<212> DNA

<213> Homo sapiens

<400> 356
cctcggaatt ccctttcagc tccagcttta cccacatcag ctgctagacg ggtacgggca 60
aaatcaagag ggtacacaaa acacagggat gtggcccctg cggcaccacc cgatgccaga 120
ttccctgcaa agtagcgcca aaactgggtt ctcttgtcca caccacccag gaagatctgc 180
ttgtatttat ctttgaaggc gaagttaaga gcctgggtgg ggaagtatct gatgacattg 240
g 241

<210> 357
<211> 234
<212> DNA
<213> Homo sapiens

<400> 357
ccaccagcag gaatgcagcg gattcctctg tcccaagtgc tcccagaagg caggattctg 60
aagaccactc cagcgatatg ttcaactatg aagaatactg caccgccaac gcagtcactg 120
ggccttgccg tgcatccttc ccacgctggt actttgacgt ggagaggaac tcctgcaata 180
acttcatcta tggaggctgc cggggcaata agaacagcta ccgctctgag gagg 234

<210> 358
<211> 615
<212> DNA
<213> Homo sapiens

<400> 358
tccccccccc ccccaaaaaa aagccatccc cccgctctgc cccgtcgcac attcggcccc 60
cgcgactcgg ccagagcggc gctggcagag gagtgtccgg caggagggcc aacgcccgct 120
gttcggtttg cgacacgcag cagggaggtg ggcggcagcg tcgccggctt ccagacacca 180
atgggaatcc caatggggaa gtcgatgctg gtgcttctca ccttcttggc cttcgcctcg 240
tgctgcattg ctgcttaccg ccccagtgag accctgtgcg gcggggagct ggtggacacc 300
ctccagttcg tctgtgggga ccgcggcttc tacttcagca ggcccgcaag ccgtgtgagc 360
cgtcgcagcc gtggcatcgt tgaggagtgc tgtttccgca gctgtgacct ggccctcctg 420
gagacgtact gtgctacccc cgccaagtcc gagagggacg tgtcgacccc tccgaccgtg 480
cttccggaca acttccccag ataccccgtg ggcaagttct tccaatatga cacctggaag 540
cagtccaccc agcgcctgcg caggggcctt gcctgccctc ctgcgtgccc gcggggtcac 600
gtgctcgcca aggag 615

<210> 359
<211> 201
<212> DNA
<213> Homo sapiens

<400> 359
ccaaactggg agaagatggc gtacagggac ttttttagct ctgcagtggg ggaaaaaaag 60
acctttgagc tccccttttt agaagaagcg cagcccaaat agaagaccat cagccaattg 120
atggggtacc cagaatgcct cgggtccata atctctcaag acctaacaat cacagaagca 180
gccaggaacc actcaagatg g 201

<210> 360
<211> 419
<212> DNA
<213> Homo sapiens

<400> 360
ctggtaggga gcaattctat tatttggcat tgcatggctg ggttgaatta aaacagggag 60
tgagaacagg tgagtctaga agtccaactc tgaaaaggac cactgtacat ttgaacacac 120
ggctgtgtta aagatgctgc taatgtcagt cactgggtgc actaaaggat ctcttatttt 180

atgtaaaacg ttgggattga caagatagat ctgatactct gttaagttac cctctgaagc 240
tacttcttgt gaaatactaa tgacagcatc atcctgccaa gcgaaagagg caggcataag 300
caaggacaaa ttaaaagggg gtaagagcct tatcatgatg aggagtcttg ttttgacatc 360
ttgggaaaag ctgtccatag tgtgaagtcg tcaatttctc accatggttt gcagtttgc 419

<210> 361
<211> 792
<212> DNA
<213> Homo sapiens

<400> 361
gcgtccctct gcctgcccac tcagtggcaa cacccgggag ctgttttgtc ctttgtggag 60
cctcagcagt tccctctttc agaactcact gccaagagcc ctgaacagga gccaccatgc 120
agtgcttcag cttcattaag accatgatga tcctcttcaa tttgctcatc tttctgtgtg 180
gtgcagccct gttggcagtg ggcatctggg tgtcaatcga tggggcatcc tttctgaaga 240
tcttcgggcc actgtcgtcc agtgccatgc agtttgtcaa cgtgggctac ttcctcatcg 300
cagccggcgt tgtggtcttt gctcttggtt tcctgggctg ctatggtgct aagactgaga 360
gcaagtgtgc cctcgtgacg ttcttcttca tcctcctcct catcttcatt gctgaggttg 420
cagctgctgt ggtcgccttg gtgtacacca caatggctga gcacttcctg acgttgctgg 480
tagtgcctgc catcaagaaa gattatggtt cccaggaaga cttcactcaa gtgtggaaca 540
ccaccatgaa agggctcaag tgctgtggct tcaccaacta tacggatttt gaggactcac 600
cctacttcaa agagaacagt gcctttcccc cattctgttg caatgacaac gtcaccaaca 660
cagccaatga aacctgcacc aagcaaaagg ctcacgacca aaaagtagag ggttgcttca 720
atcagctttt gtatgacatc cgaactaatg cagtcaccgt gggtggtgtg gcagctggaa 780
ttgggggcct cg 792

<210> 362
<211> 141
<212> DNA
<213> Homo sapiens

<400> 362
aaaggagttg gaggagaggg agggggagga catggcacca ttccagaaac cagcattgtt 60
acaacaccat agccagtata tttagtttgg cttttcctaa catagaaatc ttcaaagctg 120
gggaagtgga aataaagttt t 141

<210> 363
<211> 219
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(219)
<223> n = A,T,C or G

<400> 363
gctggcagag gagtgtccgg caggagggcc aacgcccgct gttcggtttg cgacacgcan 60
cagggaggtg ggcggcagcg tcgccggctt ccagacacca atgggaatcc caatggggaa 120
gtcgatgctg gngcttctca ccttcttggc cttcgcctcg tgctgcattg ctgcttaccg 180
ccccaatgag accctgtgcg gnggggagct ggnggacac 219

<210> 364
<211> 268
<212> DNA
<213> Homo sapiens

<220>

<221> misc_feature
<222> (1)...(268)
<223> n = A,T,C or G

<400> 364
naccactcgc tccaccttct ccaccaacta ccggnccctg ggctctgncc aggcncccag 60
ctacngcgcc cggncggtca ncagcncggc cagcgtctat gcaggcgctg ggggctctgg 120
ntcccggatc tccgtgtccc gntccaccan cttnangggc ggtatggggn ccgggggccn 180
ggccaccggg atagccnggg gnntggcagg antgggaggc ntccagaacg agaagganac 240
catgcagaga ctgaacgacc gcctggcc 268

<210> 365
<211> 151
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(151)
<223> n = A,T,C or G

<400> 365
aaggggcttt tagaaatatt taaggacaac ataaggtatt aatattggaa aaaaactgta 60
catattttca agcacaacnc tgaaatattg caacagngtt taactgaatt gttttacctg 120
cccgggcggc cgctcgaaag ggcgaattcc a 151

<210> 366
<211> 304
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(304)
<223> n = A,T,C or G

<400> 366
gccagtgtga tggatatctg cagaattcgc ccttagcgtg gtcgcggccg aggtaaacta 60
canagggttt tccagctatt atttccttta gtttctaaaa gtaacgactt atattaatgt 120
tttataaaag atagtgatga aaaaaaggta atgctgaaat aaaggcgctt ttagaaatat 180
ttaaggacaa cataaggtat taatattgga aaaaaactgt acatattttc aagcacaaca 240
ctgaaatatt gcagcagtgt ttaactgaat tgttttacct gcccgggcgg ccgctcgaaa 300
gggc 304

<210> 367
<211> 501
<212> DNA
<213> Homo sapiens

<400> 367
caggtccttc gagaagatcc ctagtgagac tttgaaccgt atcctgggcg acccagaagc 60
cctgagagac ctgctgaaca accacatctt gaagttagct atgtgtgctg aagccatcgt 120
tgcggggctg tctgtagaga ccctggaggg cacgacactg gaggtgggct gcagcgggga 180
catgctcact atcaacggga aggcgatcat ctccaataaa gacatcctag ccaccaacgg 240
ggtgatccac tacattgatg agctactcat cccagactca gccaagacac tatttgaatt 300
ggctgcagag tctgatgtgt ccacagccat tgacctttc agacaagccg gcctcggcaa 360
tcatctctct ggaagtgagc ggttgaccct cctggctccc ctgaattctg tattcaaaga 420
tggaacccct ccaattgatg cccatacaag gaatttgctt cggaaccaca taattaaaga 480

ccagctggcc tctaagtatc t 501

<210> 368
<211> 581
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(581)
<223> n = A,T,C or G

<400> 368
cgtcagtgcg atggatatct gcncaattcg ccctttgagc ggccgcccgg gcaggtcctt 60
cgataagatc cctagtgaga ctttgaaccg tatcctgggc gacccagaag ccctgagaga 120
cctgctgaac aaccacatct tgaagttagc tatgtgtgct gaagccatcg ttgcggggct 180
gtctgtagag accctggagg gcacgacact ggaggtgggc tgcagcgggg acatgctcac 240
tatcaacggg aaggcgatca tctccaataa agacatccta gccaccaacg gggtgatcca 300
ctacattgat gagctactca tcccagactc agccaagaca ctatttgaat tggctgcaga 360
gtctgatgtg tccacagcca ttgacctttt cagacaagcc ggctcggcaa tcatctctct 420
ggaagtgagc ggttgaccct cctggctccc ctgaattctg tattcaaaga tggaacccct 480
ccaattgatg cccatacaag gaatttgctt cggaaccaca taattaaaga ccagctggcc 540
tctaagtatc tgtccatgga cagaccctgg gaaactctgg g 581

<210> 369
<211> 381
<212> DNA
<213> Homo sapiens

<400> 369
aattgtgttt taattgtaaa aatggcaggg ggtggaatta ttactctata cattcaacag 60
agactgaata gatatgaaag ctgatttttt ttaattacca tgcttcacaa tgttaagtta 120
tatggggagc aacagcaaac aggtgctaat ttgttttgga tatagtataa gcagtgtctg 180
tgttttgaaa gaatagaaca cagtttgtag tgccactgtt gttttggggg ggctttttc 240
ttttcggaaa tcttaaacct taagatacta aggacgttgt tttggttgta ctttggaatt 300
cttagtcaca aaatatattt tgtttacaaa aatttctgta aaacaggtta taacagtgtt 360
tacctcggcc gcgaccacgc t 381

<210> 370
<211> 501
<212> DNA
<213> Homo sapiens

<400> 370
gtaaagtaca ttatgagaac aacagccctt tcctgaccat caccagcatg acccgagtca 60
ttgaagtctc tcactggggt aatattgctg tggaagaaaa tgtggactta aagcacacag 120
gagctgtgct taaggggcct ttctcacgct atgattacca gagacagcca gatagtggaa 180
tatcctccat ccgttctttt aagaccatcc ttcctgctgc tgcccaggat gtttattacc 240
gggatgagat tggcaatgtt tctaccagcc acctccttat tttggatgac tctgtagaga 300
tggaaatccg gcctcgcttc cctctctttg gcgggtggaa gacccattac atcgttggct 360
acaacctccc aagctatgag tacctctata atttgggtga ccagtatgca ctgaagatga 420
ggtttgtgga ccatgtgttt gatgaacaag tgatagattc tctgactgtg aagatcatcc 480
tgcctgaagg agccaagaac a 501

<210> 371
<211> 439
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(439)
<223> n = A,T,C or G

<400> 371
gtctgtcctc tgttactcan atacagnncc aaaactaagc gattatataa gcacatccat 60
attttagggc tactctaagt taaaaacctt ttctcttgtt tcagagttat ttacatcaaa 120
ttaagacatt tacaaattgt tcatagtata caatagccca aatatgattt tcacctatgc 180
tgtgtaaaga agttaagcat tcgtaagttt gtctaataaa ttcagtgcac ttttttccat 240
aacacgagct attctaaatg ttttacattt ctttcagtgc atatttccaa attcattaaa 300
cagaatgaaa tcaatgttat taaatggcta tatcataata ttcaagcata ttatggaatc 360
tataccacag tgggattcac gtcaatacta taattcactc tagaaaaaca tcacaggcac 420
acacaaaata aagaacaaa 439

<210> 372
<211> 162
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(162)
<223> n = A,T,C or G

<400> 372
taatnaaatt cattgcactt ttttccataa cacgagctat cctaagtgtt gtacatttct 60
ttcagtgcat atttncaaat tcattaaaca caatgaaatc aatgntatta aatggcnata 120
tcataanatt caancatatt atggaatgta taccacaccg gg 162

<210> 373
<211> 301
<212> DNA
<213> Homo sapiens

<400> 373
ggtccgttga ctattctcta caaaccacaa agacattgga acactatacc tattattcgg 60
cgcatgagct ggagtcctag gcacagctct aagcctcctt attcgagccg agctgggcca 120
gccaggcaac cttctaggta acgaccacat ctacaacgtt atcgtcacag cccatgcatt 180
tgtaataatc ttcttcatag taatacccat cataatcgga ggctttggca aacctgcccg 240
ggcggccaag ggcgaattcc agcacactgg cggccgttac tagtggatcc gagctcggta 300
c 301

<210> 374
<211> 471
<212> DNA
<213> Homo sapiens

<400> 374
catgtccctg atctcagtga ggtcctcctt ggtgaacaca aagcccacat tcccccggat 60
atgaggcagc agtttctcca gagctgggtt gttttccagg tgccctcgga tggccttgcg 120
catcatggtg ttcttgccca tcagcaccac agccttcccg cgaagggaca tgcggatctg 180
ctgcatctgc ttggagccca cattgtctgc tcccacaatg aaacatttcg gataatcatc 240
caatagttgg atgatcttaa ggaagtagtt ggacttccag gtcgccctgt cttccctggg 300
catcacggcg gtgcgtcagg gattgccacg cagggtttac ctcggccgcg accacgctaa 360
gggcgaattc cagcacactg gcggccgtta ctagtggatc cgagctcggt accaagcttg 420
gcgtaatcat ggtcatagct gtttcctgtg tgaaattgtt atccgctcac a 471

<210> 375
<211> 287
<212> DNA
<213> Homo sapiens

<400> 375
tatgtagcat ttagcaccac tggctcaatg cgctcaccta ggtgagagtg tgaccaaatc 60
ttaaagcatt agtgctatta tcagttacca ccatttgggg cttttatcct tcatgggtta 120
tgatgttctc ctgatgacac atttctctga gttttgtaat tccagccaaa gagagaccat 180
tcactatttg atggctggct gcatgcagac atttacctcg gccgcgacca cgctaagggc 240
gaattccagc acactggcgg ccgttactag tggatccgag ctcggta 287

<210> 376
<211> 309
<212> DNA
<213> Homo sapiens

<400> 376
tccaaccaag agtgtgctcc agatgtgttt gggccctacc tggcacagag tcctgctcct 60
gggaaaggaa aggaccacag caaacaccat tctttttgcc gtacttccta gaagcactgg 120
aagaggactg gtgatggtgg agggtgagag ggtgccgttt cctgctccag ctccagacct 180
tgtctgcaga aaacatctgc agtgcagcaa atccatgtcc agccaggcaa ccagacctgc 240
ccgggcggcc gcccgaaagg gcgaattcca gcacactggc ggccgttact agtggatccg 300
agctcggta 309

<210> 377
<211> 490
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(490)
<223> n = A,T,C or G

<400> 377
gtaccacatc gaggtgaacc gtgttcctgc tggcaactgg gttctgattg aaggtgttga 60
tcaaccaatt gtgaanacag caaccataac cgaaccccga ggcaatgagg aggctcanat 120
tttccgaccc ttgaagttca ataccacatc tgttatcaag attgctgtgg agccagtcaa 180
cccctcagag ctgcccaaga tgcttgatgg cctgcgcaag gtcaacaaga gctatccatc 240
cctcaccacc aaggtggagg agtctggcaa gcatgtgatc ctgggcactg gggagctcta 300
cctggactgt gtgatgcatg atttgcggaa gatgtactca nagatagaca tcaaggtggc 360
tgacccagtt gtcacgtttt gtgagacggt ggtggaaaca tcctccctca agtgctttgc 420
tgaaacgcct aataagaaga acaagatcac catgattgct gagcctcttg agaagggcct 480
ggcagaggac 490

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
20 December 2001 (20.12.2001)

PCT

(10) International Publication Number
WO 01/096389 A3

(51) International Patent Classification[7]: C12N 15/12, 15/11, 15/62, 5/06, C07K 14/47, 14/81, 16/18, C12Q 1/68, G01N 33/50, A61K 38/17, 31/7088

(21) International Application Number: PCT/US01/18574

(22) International Filing Date: 7 June 2001 (07.06.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/210,667 9 June 2000 (09.06.2000) US
60/252,614 22 November 2000 (22.11.2000) US

(71) Applicant *(for all designated States except US)*: **CORIXA CORPORATION** [US/US]; 1124 Columbia Street, Suite 200, Seattle, WA 98104 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **MEAGHER, Madeleine, Joy** [US/US]; 507 N.E. 71st, #1, Seattle, WA 98115 (US). **KING, Gordon, E.** [US/US]; 1530 N.W. 52nd, #304, Seattle, WA 98107 (US). **XU, Jiangchun** [US/US]; 15805 S.E. 43rd Place, Bellevue, WA 98006 (US). **SECRIST, Heather** [US/US]; 3844 35th Avenue W., Seattle, WA 98199 (US).

(74) Agents: **POTTER, Jane, E., R.**; Seed Intellectual Property Law Group PLLC, Suite 6300, 701 Fifth Avenue, Seattle, WA 98104-7092 et al. (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*

(88) Date of publication of the international search report:
19 June 2003

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 01/096389 A3

(54) Title: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS OF COLON CANCER

(57) Abstract: Compositions and methods for the therapy and diagnosis of cancer, such as colon cancer, are disclosed. Compositions may comprise one or more colon tumor proteins, immunogenic portions thereof, or polynucleotides that encode such portions. Alternatively, a therapeutic composition may comprise an antigen presenting cell that expresses a colon tumor protein, or a T cell that is specific for cells expressing such a protein. Such compositions may be used, for example, for the prevention and treatment of diseases such as colon cancer. Diagnostic methods based on detecting a colon tumor protein, or mRNA encoding such a protein, in a sample are also provided.

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 01/18574

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C12N15/12 C12N15/11 C12N15/62 C12N5/06 C07K14/47
C07K14/81 C07K16/18 C12Q1/68 G01N33/50 A61K38/17
A61K31/7088

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 C07K C12N C12Q G01N A61K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)
BIOSIS, EPO-Internal, PAJ, WPI Data, SEQUENCE SEARCH

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 99 47674 A (CORIXA CORP) 23 September 1999 (1999-09-23) see SEQ ID NO: 75 (Fig. 30) page 1 -page 32; claims 12-16,24-27,31-45; examples 1,4 ---	1-17
X	DATABASE EM_EST [Online] EBI Hinxton, UK; AC/ID AI831499, 13 July 1999 (1999-07-13) NCI-CGAP: "Homo sapiens cDNA clone IMAGE:2406189" XP002192540 abstract --- -/--	1,3,4,8, 11,15

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
14 March 2002	06. 06. 2002

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Oderwald, H

Form PCT/ISA/210 (second sheet) (July 1992)

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	PENNACCHIO LEN A ET AL: "Mutations in the gene encoding cystatin B in progressive myoclonus epilepsy (EPM1)." SCIENCE (WASHINGTON D C), vol. 271, no. 5256, 1996, pages 1731-1734, XP002192539 ISSN: 0036-8075 the whole document	1-4,8, 11,15
A	J-M FRIGERIO ET AL: "Analysis of 2166 clones from a human colorectal cancer cDNA library by partial sequencing" HUMAN MOLECULAR GENETICS, OXFORD UNIVERSITY PRESS, SURREY, GB, vol. 4, no. 1, 1995, pages 37-43, XP002111970 ISSN: 0964-6906 the whole document	
A	YEATMAN T J AND MAO W: "Identification of a differentially-expressed message associated with colon cancer liver metastasis using an improved method of differential display" NUCLEIC ACIDS RESEARCH, OXFORD UNIVERSITY PRESS, SURREY, GB, vol. 23, no. 19, 1995, pages 4007-4008, XP002099962 ISSN: 0305-1048 the whole document	
A	CHAN E-C ET AL: "Identification of novel genes that are differentially expressed in human colorectal carcinoma" BIOCHIMICA ET BIOPHYSICA ACTA, AMSTERDAM, NL, vol. 1407, no. 3, 30 September 1998 (1998-09-30), pages 200-204, XP000910494 ISSN: 0006-3002 the whole document	

-/--

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	MEAGHER M J ET AL.: "Identification of differentially expressed genes in colon tumors using cDNA subtraction and microarray analysis" PROCEEDINGS OF THE AMERICAN ASSOCIATION FOR CANCER RESEARCH ANNUAL MEETING, vol. 41, March 2000 (2000-03), page 173 XP001064177 91st Annual Meeting of the American Association for Cancer Research;San Francisco, California, USA; April 01-05, 2000, March, 2000 ISSN: 0197-016X see abstract #1109 abstract	
E,L	WO 02 02623 A (WANTANABE YOSHIHIRO ;CORIXA CORP (US); WANG TONGTONG (US); CARTER) 10 January 2002 (2002-01-10) L: priority see SEQ ID NO: 76 (Fig. 24/25) page 1 -page 5; claims 1-17; examples 1,4 page 23 -page 113	1-17
P,X	WO 00 55351 A (HUMAN GENOME SCIENCES INC ;ROSEN CRAIG A (US); RUBEN STEVEN M (US)) 21 September 2000 (2000-09-21) see SEQ ID NO: 357 and 1130 (pp. 346, 347, 1247 and 1248) page 1 -page 3; claims 1-23; example 33	1-17
P,X	WO 01 36674 A (PASKINS LYNN DORA ;BULL JOHN HENRY (GB); ELLISON GILLIAN (GB); AST) 25 May 2001 (2001-05-25) see SDEQ ID NO: 26 (Fig. 21) page 1 -page 6; claims 1-13; examples 1-7; table 1 page 12 -page 32	1-9, 11-16

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 01/18574

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. [X] Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

Although claims 9 (in so far as in vivo methods are concerned), 12, 13 and 17 are directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.

2. [ ] Claims Nos.:
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:

3. [ ] Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

1. [ ] As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. [ ] As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. [ ] As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. [X] No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

1-17 (all partially)

Remark on Protest

[ ] The additional search fees were accompanied by the applicant's protest.

[ ] No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

Invention 1: claims 1-17 partially

An isolated polynucleotide comprising SEQ ID NO: 1, an isolated polypeptide encoded by said polynucleotide. An expression vector, a host cell, an isolated antibody, a fusion protein, an oligonucleotide, a method for stimulating or expanding T cells, an isolated T cell population, a diagnostic kit, methods for stimulating an immune response, for detecting and determining the presence of cancer, for the treatment of colon cancer and for the treatment of colon cancer comprising said polynucleotide and polypeptide.

Invention 2-377: claims 1-17 partially

same as invention 1 but comprising the nucleotide sequence in the order as given in claim 1 (invention 2 is limited to SEQ ID NO: 2 and invention 377 is limited to SEQ ID NO: 377).

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No
PCT/US 01/18574

Patent document cited in search report		Publication date	Patent family member(s)		Publication date
WO 9947674	A	23-09-1999	US	6210883 B1	03-04-2001
			US	6312695 B1	06-11-2001
			AU	3094999 A	11-10-1999
			BR	9908823 A	21-11-2000
			CA	2323093 A1	23-09-1999
			CN	1299414 T	13-06-2001
			EP	1064372 A2	03-01-2001
			NO	20004631 A	15-11-2000
			WO	9947674 A2	23-09-1999
			US	2002052329 A1	02-05-2002
WO 0202623	A	10-01-2002	AU	7312701 A	14-01-2002
			WO	0202623 A2	10-01-2002
WO 0055351	A	21-09-2000	AU	3395900 A	04-10-2000
			AU	3617600 A	04-10-2000
			AU	3617700 A	04-10-2000
			AU	3619400 A	04-10-2000
			AU	3619500 A	04-10-2000
			AU	3869400 A	04-10-2000
			EP	1168917 A2	09-01-2002
			EP	1165588 A1	02-01-2002
			EP	1169469 A1	09-01-2002
			EP	1165589 A1	02-01-2002
			EP	1159420 A1	05-12-2001
			EP	1163358 A1	19-12-2001
			WO	0055173 A1	21-09-2000
			WO	0055350 A1	21-09-2000
			WO	0055351 A1	21-09-2000
			WO	0055180 A2	21-09-2000
			WO	0055174 A1	21-09-2000
			WO	0055320 A1	21-09-2000
			US	2002039764 A1	04-04-2002
			US	2002055627 A1	09-05-2002
			US	2002052308 A1	02-05-2002
			US	2002044941 A1	18-04-2002
WO 0136674	A	25-05-2001	AU	1163101 A	30-05-2001
			WO	0136674 A2	25-05-2001